PTO 21 MAR 2005

PCT/DK 03/00609
10/528563

REC'D 27 OCT 2003
WIPO PCT

Kongeriget Danmark

BEST AVAILABLE COPY

Patent application No.: PA 2002 01391

Date of filing: 20 September 2002

Applicant:
(Name and address)

Rigshospitalet
Blegdamsvej 9
DK-2100 København Ø
Denmark

Title: Methods for increasing the production of a recombinant polypeptide from a host cell

IPC: -

This is to certify that the attached documents are exact copies of the above mentioned patent application as originally filed.

Patent- og Varemærkestyrelsen · Kongeriget Danmark

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

Patent- og Varemærkestyrelsen
Økonomi- og Erhvervsministeriet

8 October 2003

Bo Z. Tidemann

PATENT- OG VAREMÆRKESTYRELSEN

Modtaget PVS
20 SEP. 2002

Methods for increasing the production of a recombinant polypeptide from a host cell

Field of the Invention

The present invention relates to methods for enhancing the production of a polypeptide from a cell by disrupting the synthesis or activity of a pitrilysin metalloprotease. In particular, the present invention relates to methods for enhancing the secretion of recombinant polypeptides from host cells such as, but not limited to, yeast and bacterial cells.

Background of the Invention

Cholecystokinin (CCK) is a vertebrate neuroendocrine peptide hormone that is expressed in both gut and brain tissues. The maturation of bioactive CCK peptides depends on post-translational tyrosine sulfation, endoproteolytic cleavages, exoproteolytic trimmings and carboxyterminal amidation. The endoproteolytic processing of the N-terminus varies with CCK-83, -58, -39, -33, -22, -8 and -5 being identified. Most of the CCK peptides are synthesized after cleavage at a single Arg residue, however, CCK-22 requires processing after a single Lys residue.

Many recombinant polypeptides have been expressed in yeast as a fusion protein to the *Saccharomyces cerevisiae* α-factor prepro-peptide to direct secretion through the secretory pathway. The best characterized yeast protease is the serine endoprotease, Kex2p (Fuller et al., 1989) which is involved in maturation of the α-mating pheromone and of killer toxin (Julius et al., 1984). Another yeast protease is Yps1p belonging to the yapsin family of glycosyl-phophatidylinositol (GPI)-anchored aspartyl proteases, which is able to rescue mating deficiency when overexpressed in a *kex2* mutant (Egel-Mitani et al., 1990). Expression of foreign proteins have shown that Yps1p and Yps2p contain endoprotease activity.

The use of host cells for the expression of recombinant polypeptides has greatly simplified the production of large quantities of commercially valuable polypeptides, which otherwise are obtainable only by purification from their native sources. There is a varied selection of expression systems currently available from which to choose for the production of any given polypeptide, including eubacterial and eucaryotic hosts. One important factor in the selection of an appropriate expression system is the ability of the host cell to produce adequate yields of the polypeptide. However, a problem frequently encountered is the high level of proteolytic enzymes produced by a given

host cell or in the culture medium. Accordingly, there is a need for further methods which enhance the production of a recombinant polypeptide from a host cell.

Metalloproteases are the most diverse of the four main types of protease, with more than 30 families identified to date. In these enzymes, a divalent cation, usually zinc, activates the water molecule. The metal ion is held in place by amino acid ligands, usually three in number. The known metal ligands are His, Glu, Asp or Lys and at least one other residue is required for catalysis, which may play an electrophillic role. Of the known metalloproteases, around half contain an HEXXH motif, which has been shown in crystallographic studies to form part of the metal-binding site.

A number of proteases dependent on divalent cations for their activity have been shown, to belong to a single family, peptidase M16. Included are insulinase, mitochondrial processing protease, pitrilysin, nardilysin and a number of bacterial proteins. These proteins do not share many regions of sequence similarity; the most noticeable is in the N-terminal section. This region includes a conserved histidine followed, two residues later by a glutamate and another histidine. In pitrilysin, it has been shown that this H-X-X-E-H motif is involved in enzymatic activity; the two histidines bind zinc and the glutamate is necessary for catalytic activity. The X can be any amino acid. Non active members of this family have lost from one to three of these active site residues.

It has previous been suggested that one could provide host cells and methods of producing proteins by expressing significantly reduced levels of a genetical modification in order to express significantly reduced levels of a metalloprotease containing an HEXXH motif in a filamentous fungal host cell, in e.g. US 5,861,280. However, metalloprotease which can be reduced by a genetical modification in order to express significantly reduced levels of said metalloprotease in a non-filamentous fungal host cell and other cells containing an motif other than HEXXH has never been described.

Summary of the Invention

Whilst investigating the role various proteases play in processing proCCK in recombinant yeast, the present inventors surprisingly noted that the deletion/disruption of *CYM1* enhanced recombinant polypeptide production and secretion. Furthermore, the present inventors have found that Cym1p belongs to a family of metalloproteases, the activity of which can be down-regulated to enhance the levels of recombinant polypeptide produced from a host cell.

Accordingly, in a first aspect the present invention provides a host cell comprising a nucleic acid sequence encoding a recombinant polypeptide in which the production of a

naturally occurring metalloprotease comprising a sequence provided in SEQ ID NO:1 has been reduced or inhibited by genetic manipulation.

The host cell can be any cell which, in its native state, possesses the metalloprotease. Accordingly, the host cell can be a eukaryotic or prokaryotic cell. Examples of preferred eukaryotic cells include, but are not limited to, mammalian cells, plants cells and fungal cells. In a preferred embodiment, the host cell is a yeast cell. More preferably, the yeast cell is selected from, but not limited to, the group consisting of: Saccharomyces sp. such as Saccharomyces cerevisiae and Saccharomyces kluyveri, Schizosaccharomyces sp. such as Schizosaccharomyces pombe, Kluyveromyces lactis, Hansenula polymorpha, Pichia pastoris, Pichia methanolica, Pichia kluyveri, Yarrowia lipolytica, Candida utilis, Candida cacaoi, and Geotrichum fermentans.

The metalloprotease is a member of the pitrilysin subfamily of proteases, characterized by comprising the sequence HXXEH (SEQ ID NO:1) where X is any amino acid. In a preferred embodiment, the metalloprotease comprises a consensus sequence provided in SEQ ID NO:2. Even more preferably, the metalloprotease comprises a consensus sequence provided in SEQ ID NO:3. In addition, it is preferred that the metalloprotease comprises SEQ ID NO:1 and a glutamic acid residue between 70 and 80 amino acids C-terminal of the second His residue. Further, it is preferred that the metalloprotease comprises a sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2 and SEQ ID NO:3 as well as a sequence selected from the group of:

i) any one of group consisting of SEQ ID NO's 4 to 15, and
ii) a sequence which is at least 80% identical to any one of SEQ ID NO's 4 to 15.

More preferably, the metalloprotease comprises SEQ ID NO:1 or SEQ ID NO:2 or SEQ ID NO:3 as well as a sequence selected from the group of:

i) any one of SEQ ID NO's 4 or 5, and
ii) a sequence which is at least 80% identical to any one of SEQ ID NO's 4 or 5.

Preferably, the metalloprotease comprises a sequence which is at least 85% identical, such as at least 90% identical, such as at least 95% identical, and such as at least 99% identical to any one of SEQ ID NO's 4 to 15.

In a particularly preferred embodiment, the metalloprotease comprises a sequence as provided in SEQ ID NO:4, or a sequence at least 80% identical, such as at least 90%, such as at least 95% and such as 99% identical, thereto.

The host cell can be genetically manipulated by any means known in the art as long as the production of the metalloprotease is reduced or inhibited when compared to a parental host cell which has not been genetically manipulated. Such means of genetically manipulating the host cell include, but are not limited to; gene knockout, gene disruption, random or site directed mutagenesis, introduction of dominant-negative metalloproteases, RNA interference (RNAi) using dsRNA, catalytic nucleic acids (such as ribozymes and DNAzymes), and antisense nucleic acids. Preferably, the genetic manipulation acts directly upon the gene encoding the metalloprotease, the mRNA transcribed from the gene, or produces a protein that alters the activity of the metalloprotease such as a dominant negative mutant which competes with the metalloprotease for binding to a substrate but does not, for example, possess catalytic activity. However, the host cell may be genetically manipulated such that it indirectly affects the production or activity of the metalloprotease. For instance, the genetic manipulation can target a transcription factor involved in transcribing the mRNA encoded by the metalloprotease gene, thus at least reducing the levels of metalloprotease produced by the manipulated host cell.

Furthermore, the host cell may be further genetically manipulated such that it lacks at least one other naturally occurring protease of the host cell or has reduced activity for at least one other naturally occurring protease of the host cell. The protease can be any enzyme of which the inhibition increases the production of a recombinant polypeptide produced by the host cell. The protease can either be an endopeptidase, an aminopeptidase or a carboxypeptidase. Preferred proteases include, serine proteases, aspartyl proteases, cysteine proteases and other metalloproteases.

In one embodiment, the host cell is a yeast cell and the other naturally occurring protease(s) is at least one protease selected from the group consisting of; KEX2, YPS1 (previously known as YAP3), YPS2 (previously known as MKC7), YPS3, YPS6, YPS7, BAR1, STE13, KEX1, PRC1, PEP4 (also known as PRA1), APE1, APE2, APE3 and CPS1. Preferably, the host cell is a yeast cell and KEX2 production has been disrupted.

The recombinant polypeptide can be any desired polypeptide which is capable of being produced in the host cell. The recombinant polypeptide can comprise a naturally occurring sequence or have been produced by the intervention of man (e.g. a mutant or truncation of a naturally occurring protein, or a fusion between at least two different polypeptides). Typically, the recombinant polypeptide will be of commercial value, for example in the treatment of diseases.

The recombinant polypeptide can be any size. Typically, the recombinant polypeptide will range in size from about 30 amino acids to about 4,500 amino acids. In one embodiment, the recombinant polypeptide is between about 30 to about 200 amino acids in length.

In at least some host cell expression systems for producing recombinant polypeptides, it is desirable to direct the recombinant polypeptide to be secreted from the host cell. Thus, in a preferred embodiment the nucleic acid comprises a sequence which encodes a signal for directing the recombinant polypeptide to be secreted from the host cell. Preferably, the signal is an N-terminal hydrophobic signal sequence. Such N-terminal hydrophobic signal sequences are known in the art, and include, for example but not limited to, the leader sequence originating from the fungal amyloglucosidase (AG) gene such as glaA - both 18 and 24 amino acid versions e.g. from *Aspergillus sp.*, the α-factor gene such as yeasts e.g. from *Saccharomyces sp.* and *Kluyveromyces sp.*, the P-factor of *Schizosaccharomyces sp.*, and the α-amylase gene from *Bacillus sp.* In one embodiment, the recombinant polypeptide is expressed as a fusion of an N-terminal hydrophobic signal sequence and a second polypeptide sequence encoding the recombinant polypeptide which is from a different source than the signal sequence.

The nucleic acid encoding the recombinant polypeptide can be provided to the host cell using any technique known in the art. In one embodiment, the nucleic acid is inserted into the genome of the host cell using, for example, homologous recombination based techniques. In another embodiment, the nucleic acid is transfected or transformed into the host cell in an expression vector which remains extrachromosomal. For example, the expression vector can be a plasmid or a virus. Further, it is preferred that the vector comprises a selectable marker which can be used to selectively propagate host cells comprising the vector. Such selectable markers and the use thereof are also known in the art.

In a second aspect, the present invention provides a method of producing a recombinant polypeptide, the method comprising culturing a host cell according to the second aspect under suitable conditions such that the recombinant polypeptide is produced, and recovering the recombinant polypeptide.

Since the proteolytic action arising from the metalloprotease has been reduced or inhibited, the method of the second aspect of the invention improves the stability of the recombinant polypeptide produced by the host cell.

In a preferred embodiment, the recombinant polypeptide is secreted from the host cell. Furthermore, it is preferred that the secreted protein is recovered during exponential growth of a culture comprising the host cell.

Preferably, the quantity of the recovered recombinant polypeptide is higher than if a parental host cell was used. More preferably, the quantity of the recovered recombinant polypeptide is at 50% higher than if a parental host cell was used.

In a third aspect, the present invention provides a method of cleaving a polypeptide at a basic residue, the method comprising contacting the polypeptide, in the presence of a divalent cation, with a metalloprotease comprising a sequence selected from the group of:

i) any one of SEQ ID NO's 4 to 15, and

ii) a sequence which is at least 80% identical to any one of SEQ ID NO's 4 to 15.

In a particularly preferred embodiment, the metalloprotease comprises a sequence as provided in SEQ ID NO:4, or a sequence at least 80% identical, such as at least 90%, such as at least 95% and such as 99% identical, thereto.

Preferably, the metalloprotease cleaves the polypeptide at the C-terminal side of an amino acid, or sequence of amino acids, selected from the group consisting of; Lys, Arg, ArgArg, LysLys, ArgLys and LysArg. Accordingly, it is preferred that the polypeptide comprises Lys, Arg, ArgArg, LysLys, ArgLys or LysArg. Other sequence requirements may also be necessary for cleavage, however, these can readily be determined by routine experimentation.

Preferably, the divalent cation is selected from the group consisting of: Zn^{2+}, Co^{2+} and Mn^{2+}.

The method of the third aspect can be performed *in vivo*, within a recombinant host cell producing the metalloprotease, or *in vitro* in suitable reaction conditions. Considering the present disclosure, the skilled addressee could readily perform the method of the third aspect. An example of an *in vitro* system for cleaving a polypeptide with the defined metalloprotease is provided herein. In this instance, the polypeptide is contacted with the metalloprotease provided a crude yeast cell extract in 0.1 M NaH_2PO_4 (pH 4.5) and in the presence of 1 mM Mn^{2+} and 1 mM bestatin. In another

example, the metalloprotease can be recombinantly produced as a fusion protein with a suitable "tag", such as a His-tag, which enables easy purification of the fusion protein. Preferably, such a "tag" is removed (for example by enzymatic cleavage) before the metalloprotease is exposed to the substrate polypeptide.

In a fourth aspect, the present invention provides a method of identifying an agent that inhibits the activity of a metalloprotease comprising a sequence provided in SEQ ID NO:1, the method comprising the steps of:

(a) incubating the metalloprotease with the agent, in the presense of a divalent cation and a suitable substrate;
(b) determining the activity the metalloprotease on the substrate;
(c) comparing the activity obtained in step b) with the activity of a control sample that has not been incubated with the agent; and
(d) selecting an agent that inhibits the activity of the metalloprotease.

The substrate can be any polypeptide that can be cleaved by the metalloprotease and the cleavage event detected. One example disclosed herein is the use of CCK as a substrate, where the cleavage event is detected by the production of CCK-22. Similar assays can readily be developed for other substrates.

In a preferred embodiment of the fourth aspect, the metalloprotease comprises a sequence as provided in SEQ ID NO:4, or a sequence at least 80% identical, such as at least 90%, such as at least 95% and such as 99% identical, thereto.

In a fifth aspect, the present invention provides a method of producing a recombinant polypeptide, the method comprising culturing a host cell comprising a nucleic acid sequence encoding a recombinant polypeptide under suitable conditions such that the recombinant polypeptide is produced, and recovering the recombinant polypeptide wherein said culturing comprises the presence of an inhibitor of a metalloprotease comprising a sequence provided in SEQ ID NO:1.

Preferably, the inhibitor is identified according to a method of the fifth aspect.

As will be apparent, preferred features and characteristics of one aspect of the invention may be applicable to other aspects of the invention.

Throughout this specification the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element, integer or

step, or group of elements, integers or steps, but not the exclusion of any other element, integer or step, or group of elements, integers or steps.

The invention will hereinafter be described by way of the following non-limiting Figures and Examples.

Brief Description of the Accompanying Drawings

Figure 1. Cholecystokinin expression construct. PreproMfα1p-proCCK fusion protein with the amino acid sequences around the fusion site and of the primary cleavage sites shown. The major forms of secreted CCK with their N- and C-terminal amino acid residues are shown below.

Figure 2. CCK-22 maturation in cells and media as a function of cell growth. BJ2168 expressing preproMfα1p-proCCK fusion protein. The CCK-22 immunoreactivity was measured by RIA using Ab 89009 and total CCK content measured with Ab 89009 after tryptic cleavage. Open circles represent the fraction of secreted CCK-22, whereas the intracellular fraction of CCK-22 is presented as filled triangles. The cell growth was measured by OD_{600} (open squares). The data represents mean of two independent experiments.

Figure 3. Chromatographic analyses of normal and K→A mutated CCK secreted from BJ2168. Media from yeast transformed with pRS426 preproMfα1p-proCCK and pRS426 preproMfα1p-proCCK (K→A) were subjected to G-50 gel chromatography and the CCK immunoreactivity was measured with Ab 7270 specific for Gly extended CCK (A and C) and Ab 89009, which is specific for the N-terminus of CCK-22 (B and D).

Figure 4. *In vitro* protease assay including inhibitors and activators. The fraction of CCK-22 was calculated from the immunoreactivity using Ab 89009 divided by the total amount of mature and N-terminal extended CCK-22 measured with Ab 89009 after trypsin treatment. A, Effect of different inhibitors. B, Protease reactivation by addition of 1.2 mM divalent metal ions to extracts where the activity had been inhibited with 1 mM EDTA. The data represents mean ± SD of three independent experiments.

Figure 5. Protease reactivation by Zn^{2+} and Mn^{2+}. *In vitro* protease assays performed with cell extracts from LJY123, where the activity was inhibited with 1 mM EDTA (filled squares) and reactivated by addition of 1.2 mM Mn^{2+} (open circles) or 1.2 mM Zn^{2+} (filled circles). The activity was measured as the fraction of matured CCK-22 after 30,

60 and 120 min incubation. The fraction of CCK-22 was calculated from the immunoreactivity using Ab 89009 divided by the total amount of mature and N-terminal extended CCK-22 measured with Ab 89009 after trypsin treatment. The data are represented by mean ± SD (n=3).

Figure 6. Extracellular CCK-22 maturation by members of the yapsin family. Intact cells ability to process extracellular CCK was analysed as described under "Experimental Procedures" for BY4705 and the isogenic *yps1, yps1 yps3* and *yps1 yps2 yps3* strains. The fraction of CCK-22 was calculated from the immunoreactivity using Ab 89009 divided by the total amount of total CCK measured with Ab 89009 after trypsin treatment. The data are represented by mean ± SD (n=4). Statistics were performed using unpaired t test as described in experimental procedures (*** = P<0.001, ** = P<0.01 and * = P<0.05).

Figure 7. Increased proteolysis following *Cym1p* overexpression. *In vitro* protease assays performed with cell extracts from BJ2168 transformed with an empty pRS425 plasmid (A) and with pRS425 containing *CYM1* (B). The CCK-22 immunoreactivity was measured over time using Ab 89009 (filled squares and circles) and the total amount of mature and N-terminal extended CCK-22 measured with Ab 89009 after trypsin treatment (open squares and circles). The data represents mean ± SD of three independent experiments.

Figure 8. Effects of *KEX2* and *CYM1* deletions on proCCK secretion and CCK-22 maturation. Yeast cells transformed with the proCCK expression construct were harvested during exponential phase and the media collected. The intracellular (A) and extracellular (B) amount of total CCK was measured with Ab 89009 after trypsin treatment. The fraction of intracellular (C) and secreted (D) CCK-22 was calculated as the immunoreactivity measured with Ab 89009 before tryptic cleavage divided with the total amount of CCK measured in (A) and (B). The *kex2*, *cym1* and *kex2 cym1* strains are isogenic to BJ2168. The data are given as mean ± SD (n=4). Statistics were performed using unpaired t test (*** = P<0.001, ** = P<0.01 and * = P<0.05). The stars enclosed in brackets are comparison between the *kex2* and *kex2 cym1* strain.

Figure 9. Intracellular degradation of CCK depends on Cym1p cleavage to CCK-22. Expression of wild type CCK, preproMfα1p-proCCK, and the CCK mutant, preproMfα1p-proCCK (K→A) in BJ2168 and a *CYM1* disrupted strain isogenic to BJ2168. The cells were sedimented during exponential growth and the total amount of CCK (hatched

bars) was measured after trypsin and carboxypeptidase B treatment with Ab 7270 specific for Gly-extended CCK. The amount of mature Gly-extended (white bars), which is dependent on translocation into the secretory pathway, Kex2p and carboxypeptidase activity is measured as the immunoreactivity using Ab 7270 before tryptic cleavage and carboxypeptidase B treatment. The data are given as mean ± SD (n=3).

Figure 10. Aspartyl proteases involved in the maturation of CCK-22. Expression of wild type proCCK in BY4705 and the isogenic yapsin deletion strains of *YPS1*, *YPS2* and *YPS3*. The intra- (A) and extracellular (B) fraction of synthesised CCK-22 was measured during exponential growth. The fraction of mature CCK-22 was calculated as the immunoreactivity measured with Ab 89009 before tryptic cleavage divided with the total amount of CCK. The data are represented by mean ± SD (n=3). Statistics were performed using unpaired t test as described in experimental procedures (*** = P<0.001, ** = P<0.01 and * = P<0.05).

Figure 11. Cym1p processing C-terminally to both Lys and Arg residues. *CYM1* deletion enhance the amount of secreted CCK with more than two fold of both wild type CCK and the $Lys^{61} \rightarrow Arg^{61}$ mutant. Expression of wild type CCK, preproMfα1p-proCCK, and the CCK mutant, preproMfα1p-proCCK ($Lys^{61} \rightarrow Arg^{61}$) in BJ2168 and a *CYM1* disruptant isogenic to BJ2168. Yeast cells were harvested during exponential phase and the media collected. The intracellular (A) and extracellular (B) amount of total CCK was measured with Ab 89009 after trypsin treatment. The data are given as mean ± SD (n=3).

Figure 12. Secreted proCCK fragments identified by mass spectrometry. The CCK-numbers refer to C-terminal amidated CCK. The molecular masses are given as monoisotopic values except for * which denote average value. Strain A, vacuolar protease-deficient strain (BJ2168), and B, the isogenic strain with *KEX1 KEX2* disruptions (LJY22).

Figure 13. Model for the production of the C-terminally extended CCK (A) and GLP2 (B). Expression of these fusion peptides should be performed in a *sec61* mutant, or the pre-sequence of the α-mating factor should be removed to avoid translocation into the ER. The amino acid sequences around the fusion sites are shown. Underlined are the N- and C-terminal amino acids of the Gly-extended CCK-22 and GLP1.

Key to the Sequence Listing

SEQ ID NO:1 – Consensus sequence for pitrilysin proteases.

SEQ ID NO:2 - Consensus sequence for at least some pitrilysin proteases.

SEQ ID NO:3 - Consensus sequence for at least some pitrilysin proteases.

SEQ ID NO:4 – *Saccharomyces cerevisiae* Cym1p (Swissprot Accession No. P32898).
SEQ ID NO:5 - *Schizosaccharomyces pombe* C119.7 (Swissprot Accession No. O42908).
SEQ ID NO:6 - *Clostridium perfringens* HypA protein (Swissprot Accession No. Q46205).
SEQ ID NO:7 - *Borrelia burgdorferi* protein BB0228 (Swissprot Accession No. O51246).
SEQ ID NO:8 - *Caenorhabditis elegans* C05D11.1 protein (Swissprot Accession No. P48053).
SEQ ID NO:9 - *E. coli* protease III (Swissprot Accession No. P05458).
SEQ ID NO:10 - Human NRD convertase (Swissprot Accession No. P47245).
SEQ ID NO:11 - Human insulysin (Swissprot Accession No. P14735).
SEQ ID NO:12 - *Arabidopsis thaliana* CPE (Genbank Accession No. T03302).
SEQ ID NO:13 – Human metalloprotease I (GenBank Accession No. AAH01150).
SEQ ID NO:14 – *Bacillus subtilis* zinc protease ymxG (GenBank Accession No. Q04805).
SEQ ID NO:15 – *Mycobacterium tuberculosis* zinc protease Rv2782c (GenBank Accession No. O33324).
SEQ ID NO's 16 to 42 – Oligonucleotides.
SEQ ID NO's 43 to 52 – Sequences provided in Figure 12.
SEQ ID NO's 53 to 55 – Sequences provided in Figure 1.

Detailed Description of the Invention

The present invention provides a host cell useful for the expression of a polypeptide, said cell being genetically manipulated in order to at least produce reduced levels of a defined metalloprotease, when compared to the parental cell. The host cell will thus be able to express a protein of interest in higher quantity due to the proteolytic action of the metalloprotease has been reduced or inhibited which improves the stability of the protein of interest.

By the method of the invention, the proteolytic action of the metalloprotease has been reduced or inhibited, thereby improving the stability of the product obtained.

One embodiment of the present invention relates to a host cell useful for the expression of a protein of interest, wherein said cell has been genetically modified in order to express significantly reduced levels of a metalloprotease which is at least 80% identical to the SEQ ID NO: 4, as compared to a parental cell.

In the present context, the term "protein of interest" relates to any of the numerous naturally native occurring extremely complex substances such as but not limited to proteins, enzymes and/or antibodies that consist of amino acid residues joined by peptide bonds. It is an object of preferred embodiments of the present invention to

provide such native proteins which are products of the host cell itself and/or heterologous proteins, fusion proteins, recombinant proteins, eucaryotic proteins, prokaryotic proteins, lysosomal proteins, vacuolar proteins, precursor proteins, zymogene proteins, prepro-proteins, and secreted proteins.

In the present context, the term "host cell" relates to any cell capable of producing the protein of interest. Thus in one prefered embodiment, the host is a prokaryotic cell. In another preferred embodiment, the host cell is a eukaryotic cell, such as but not limited to a filamentous fungal cell and a non-filamentous fungal cell. Non limiting examples hereof are a strain of *Saccharomyces*, especially *Saccharomyces cerevisiae*.

All the features desricbed below relating to the method of the present invention are also applicable as embodiments relating to the host cell.

The method described in the present application relates to the production of a protein of interest in a host cell, wherein said host cell has been genetically modified in order to express significantly reduced levels of a metalloprotease which is at least 80% identical to the SEQ ID NO: 4 as compared to a parental cell, when cultured under identical conditions, comprising

a) introducing into the host cell a nucleic acid sequence encoding the protein of interest;

b) cultivating the host cell of step (a) in a suitable growth medium for production of the protein of interest and

c) isolating the protein of interst.

One embodiment of the present invention relates to a method for the production of a protein of interest in a host cell, wherein the host cell has been genetically modified by a method selected form the group comprising gene knock-out, gene disruption, random or site directed mutagenesis, introduction of dominant-negative metalloproteases, RNA interference (RNAi) using dsRNA, catalytic nucleic acids (such as ribozymes and DNAzymes), antisense nucleic acids or a combination thereof.

In a most preferred embodiment the host cell is essentially free of any metalloprotease activity.

One preferred embodiment of the present invention relates to a method for the production of a protein of interest in a host cell, in which the protein of interest is a eucaryotic protein, selected from the group comprising insulin, growth hormone, glucagon, somatostatin, inteferon, adrenocorticotropic hormones, angiotensinogen, atrial natriuretic peptides, dynorphin, endorphines, galanin, gastrin, gastrin releasing

peptides, neuropeptide Y fragments, pancreastatin, pancreatic polypeptides, secretin, vasoactive intestinal peptide, growth hormone releasing factor, melanocyte stimulating hormone, neurotensin, adrenal peptide, parathyroid hormone and related peptides, somatostatin and related peptides, brain natriuretic peptide, calcitonin, corticotropin releasing factor (CRF), cocaine amphetamine regulated transcript (CART), thymosin, urotensin, glucagon and glucagon like peptides (GLP-1 and GLP-2), somatostatin, interferon, a vascular endothelial growth factor (VEGF), platelet derived growth factor (PDGF), factor VII, factor VIII, factor V, factor IX, interleukins, urokinase, erythropoietin (EPO), chymosin, tissue plasminogen activator, CCK or serum albumin.

Another preferred embodiment of the present invention relates to a method for the production of a protein of interest in a host cell, in which the protein of interest is a protein of fungal origin, selected from the group comprising an amylolytic enzyme, an alpha-amylase, a beta-amylase, a glycoamylase, a alpha-galactosidase, a cellulytic enzyme, a lipolytic enzyme, a xylanolytic enzyme, a proteolytic enzyme, an oxidoreductase, a peroxidase, a laccase, a pectinase, or a cutinase.

A further preferred embodiment of the present invention relates to a method for the production of a protein of interest in a host cell, in which the protein of interst is a bacterial protein, selected from the group comprising an amylolytic enzyme, an alpha-amylase, a beta-amylase, a glycoamylase, a beta-galactosidase, a cellulytic enzyme, a lipolytic enzyme, a xylanolytic enzyme, a proteolytic enzyme, an oxidoreductase, a peroxidase, a laccase, a pectinase, or a cutinase.

A special embodiment of the present invention relates to a method for production of a protein of interest in a host cell, in which the protein of interest is a precursor, i.e. a zymogen, a hybrid protein, a protein obtained as a pro sequence or pre-pro sequence, or in unmaturated form.

General Molecular Biology

Unless otherwise indicated, the recombinant DNA techniques utilised in the present invention are standard procedures, well known to those skilled in the art. Such techniques are described and explained throughout the literature in sources such as, J. Perbal, A Practical Guide to Molecular Cloning, John Wiley and Sons (1984), J. Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbour Laboratory Press (1989), T.A. Brown (editor), Essential Molecular Biology: A Practical Approach, Volumes 1 and 2, IRL Press (1991), D.M. Glover and B.D. Hames (editors), DNA Cloning: A Practical Approach, Volumes 1-4, IRL Press (1995 and 1996), F.M. Ausubel et al. (Editors), Current Protocols in Molecular Biology, Greene Pub. Associates

and Wiley-Interscience (1988, including all updates until present), Methods in Enzymology. Vol 194. Guide to Yeast Genetics and Molecular Biology. (1991) Ed Gunthrie and Fink Academic Press, Methods in Microbiology Vol. 26. Yeast Gene Analysis. (1998) Ed. Brown and Tuite. Academic Press, Miller, J. H. (1992) *A Short Course in Bacterial Genetics* (Manual, L., ed), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, Johnston, J. R (1994) *Molecular Genetics of Yeast* (A Practical Approach) Oxford University Press, Oxford., and Molecular Genetics of Yeast: A Practical Approach, Ed. J.R. Johnston, IRL Press (1994) and are incorporated herein by reference.

The % identity of a polypeptide is determined by GAP (Needleman and Wunsch, 1970) analysis (GCG program) with a gap creation penalty=5, and a gap extension penalty=0.3. The query sequence is at least 15 amino acids in length, and the GAP analysis aligns the two sequences over a region of at least 15 amino acids. More preferably, the query sequence is at least 50 amino acids in length, and the GAP analysis aligns the two sequences over a region of at least 50 amino acids. Even more preferably, the query sequence is at least 100 amino acids in length and the GAP analysis aligns the two sequences over a region of at least 100 amino acids. More preferably, the query sequence is at least 250 amino acids in length and the GAP analysis aligns the two sequences over a region of at least 250 amino acids. Even more preferably, the query sequence is at least 500 amino acids in length and the GAP analysis aligns the two sequences over a region of at least 500 amino acids.

Pitrilysin Subfamily of Metalloproteases

The pitrilysin subfamily of metalloproteases is characterized by the presence of a HXXEH (SEQ ID NO:1) motif. A general review of this subfamily is provided by Rawlings and Barrett (1995). Members of this family include, but are not limited to, *S. cerevisiae* Cym1p (SEQ ID NO:4) (Swissprot Accession No. P32898), *Schizosaccharomyces pombe* C119.7 (SEQ ID NO:5) (Swissprot Accession No. O42908), *Clostridium perfringens* HypA protein (SEQ ID NO:6) (Swissprot Accession No. Q46205), *Borrelia burgdorferi* protein BB0228 (SEQ ID NO:7) (Swissprot Accession No. O51246), *Caenorhabditis elegans* C05D11.1 protein (SEQ ID NO:8) (Swissprot Accession No. P48053), *E. coli* protease III (SEQ ID NO:9) (Swissprot Accession No. P05458), human NRD convertase (SEQ ID NO:10) (Swissprot Accession No. P47245), human insulysin (SEQ ID NO:11) (Swissprot Accession No. P14735), *Arabidopsis thaliana* CPE (SEQ ID NO:12) (Genbank Accession No. T03302), human metalloprotease I (SEQ ID NO:13) (GenBank Accession

No. AAH01150), *Bacillus subtilis* zinc protease ymxG (SEQ ID NO:14) (GenBank Accession No. Q04805), and *Mycobacterium tuberculosis* zinc protease Rv2782c (SEQ ID NO:15) (GenBank Accession No. O33324). For *E. coli* protease III (SEQ ID NO:9) it has been shown that the His residues of SEQ ID NO:1, as well as Glu-169, are involved in divalent cation binding whilst the Glu residue flanked by the His residues is a catalytic residue.

A gene encoding a pitrilysin metalloprotease can readily be identified by screening by hybridization for nucleic acid sequences coding for all of, or part of, the metalloprotease, e.g. by using synthetic oligonucleotide probes, that may be prepared on the basis of a cDNA sequence, e.g. the nucleotide sequences encoding any one of the metalloproteases presented as SEQ ID NO's: 4 to 15, in accordance with standard techniques.

Genetic Manipulations

The host cell of the invention which is genetically manipulated in order to produce reduced levels of the defined metalloprotease may be modified using standard recombinant DNA technology known to the person skilled in the art. The gene sequence responsible for the production of the metalloprotease may be inactivated or eliminated entirely.

In a particular embodiment, the host cell of the invention is one genetically manipulated at the coding or regulatory regions of the metalloprotease gene. Known and useful techniques include, but are not limited to, gene knockout, gene disruption, random or site directed mutagenesis, introduction of dominant-negative metalloproteases, RNA interference (RNAi) using dsRNA, catalytic nucleic acids (such as ribozymes and DNAzymes), and antisense nucleic acids, or a combination thereof.

Mutagenesis may be performed using a suitable physical or chemical mutagenizing agent. Examples of a physical or chemical mutagenizing agent suitable for the present purpose includes ultraviolet (UV) irradiation, hydroxylamine, N-methyl-N'-nitro-N-nitrosoguanidine (MNNG), O-methyl hydroxylamine, nitrous acid, ethyl methane sulphonate (EMS), sodium bisulfite, formic acid, and nucleotide analogues. When such agents are used, the mutagenesis is typically performed by incubating the cell to be mutagenized in the presence of the mutagenizing agent of choice under suitable

conditions for the mutagenesis to take place, and selecting for mutated cells having a significantly reduced production of metalloprotease.

Genetic manipulation may also be accomplished by the introduction, substitution or removal of one or more nucleotides in the metalloprotease coding sequence or a regulatory element required for the transcription or translation thereof. Nucleotides may, for example, be inserted or removed so as to result in the introduction of a stop codon, the removal of the start codon or a change of the open reading frame. The modification or inactivation of the structural sequence or a regulatory element may be accomplished by site-directed mutagenesis or PCR generated mutagenesis in accordance with methods known in the art.

A convenient way to inactivate or reduce the metalloprotease production of a host cell is based on the principles of gene interruption. This method involves the use of a DNA sequence corresponding to the endogenous gene or gene fragment which it is desired to destroy. The DNA sequence is *in vitro* mutated to a defective gene and transformed into the host cell. By homologous recombination, the defective gene replaces the endogenous gene or gene fragment. It may be desirable that the defective gene or gene fragment encodes a marker which may be used for selection of transformants in which gene encoding the metalloprotease has been modified or destroyed.

The term "antisense" as used herein refers to nucleotide sequences which are complementary to a specific nucleic acid sequence. Antisense molecules may be produced by any method, including synthesis by ligating the gene(s) of interest in a reverse orientation to a viral promoter which permits the synthesis of a complementary strand. Once introduced into a host cell, this transcribed strand combines with natural sequences, in this instance that encoding the metalloprotease, produced by the cell to form duplexes. These duplexes then block either the further transcription or translation. In this manner, mutant phenotypes may be generated.

The term catalytic nucleic acid refers to a DNA molecule or DNA-containing molecule (also known in the art as a "deoxyribozyme") or an RNA or RNA-containing molecule (also known as a "ribozyme") which specifically recognizes a distinct substrate and catalyzes the chemical modification of this substrate. The nucleic acid bases in the catalytic nucleic acid can be bases A, C, G, T and U, as well as derivatives thereof. Derivatives of these bases are well known in the art.

Typically, the catalytic nucleic acid contains an antisense sequence for specific recognition of a target nucleic acid, and a nucleic acid cleaving enzymatic activity, also referred to herein as the "catalytic domain". The types of ribozymes that are particularly useful in this invention are the hammerhead ribozyme (Haseloff and Gerlach, 1988) and the hairpin ribozyme (Shippy et al., 1999).

Ribozymes useful for the methods of the invention, and DNA encoding the ribozymes, can be chemically synthesized using methods well known in the art. The ribozymes can also be prepared from a DNA molecule (that upon transcription, yields an RNA molecule) operably linked to an RNA polymerase promoter, e.g., the promoter for T7 RNA polymerase or SP6 RNA polymerase. When the vector also contains an RNA polymerase promoter operably linked to the DNA molecule, the ribozyme can be produced *in vitro* upon incubation with RNA polymerase and nucleotides. In a separate embodiment, the DNA can be inserted into an expression cassette or transcription cassette. After synthesis, the RNA molecule can be modified by ligation to a DNA molecule having the ability to stabilize the ribozyme and make it resistant to RNase. Alternatively, the ribozyme can be modified to the phosphothio analog for use in liposome delivery systems. This modification also renders the ribozyme resistant to endonuclease activity.

dsRNA (RNAi) is particularly useful for specifically inhibiting the production of a particular protein. This technology relies on the presence of dsRNA molecules that contain a sequence that is essentially identical to the mRNA of the gene of interest, in this case a mRNA encoding the metalloprotease. Conveniently, the dsRNA is produced in a single open reading frame in a recombinant vector or host cell, where the sense and anti-sense sequences are flanked by an unrelated sequence which enables the sense and anti-sense sequences to hybridize to form the dsRNA molecule with the unrelated sequence forming a loop structure. The design and production of suitable dsRNA molecules for genetic manipulation is well within the capacity of a person skilled in the art, particularly considering Waterhouse et al. (1998), Elbashir et al. (2001), WO 99/32619, WO 99/53050, WO 99/49029, and WO 01/34815.

Owing to the genetic manipulation, the host cell of the invention expresses significantly reduced levels of the metalloprotease. In a preferred embodiment, the level of metalloprotease expressed by the host cell is reduced more than about 25%, such as more than about 30%, such as more than about 35%, such as more than about 40%, such as more than about 45%, such as more than about 50%, such as more than about 55%,such as more than about 60%, such as more than 65about %, such as

more than about 70%, such as more than about 75%, such as more than about 80%, such as more than about 85%, such as more than about 90%, such as more than about 95%, such as more than about 98%, and such as more than about 99%.

In a most preferred embodiment, the product expressed by the host cell is essentially free of any activity of the defined metalloprotease.

In the present context, the term "essentially free" relates to a host, wherein the metalloprotease expressed by said host cell is reduced to a level, where the function of said metalloprotease has no biological significant reducing influence on the production of the protein of interest.

Protein of Interest

The terms "polypeptide", "protein" and "peptide" are used herein interchangeably and in the present context relates to any of the numerous naturally occurring extremely complex substances such as but not limited to enzymes or antibodies that consist of amino acid residues joined by peptide bonds, contain the elements carbon, hydrogen, nitrogen, oxygen, usually sulphur, and occasionally other elements such as but not limited to phosphorus or iron, that are essential constituents of all living cells, that are in nature synthesised from raw materials by plants but assimilated as separate amino acids by animals, that are both acidic and basic and usually colloidal in nature although many have been crystallised, and that are hydrolyzable by acids, alkalies, proteolytic enzymes, and putrefactive bacteria to polypeptides, to simpler peptides, and ultimately to alpha-amino acids.

As defined herein, a "recombinant polypeptide" is a protein which is not native to the host cell, or a native polypeptide in which modifications have been made to alter the native sequence, or a native protein whose expression is quantitatively altered as a result of a manipulation of a native regulatory sequence required for the expression of the native protein, such as a promoter, a ribosome binding site, etc., or other manipulation of the host cell by recombinant DNA techniques.

Owing to the absence or reduction in activity of the defined metalloprotease, at least a portion of the recombinant polypeptides expressed by the host cell may also be a precursor protein, i.e. a zymogen, a hybrid protein, a protein obtained as a pro sequence or pre-pro sequence, or in unmaturated form.

In a more specific embodiment, the recombinant polypeptide is of eucaryotic origin, such as insulin, adrenocorticotropic hormones, angiotensinogen, atrial natriuretic peptides, dynorphin, endorphines, galanin, gastrin, gastrin releasing peptides, neuropeptide Y fragments, pancreastatin, pancreatic polypeptides, secretin, vasoactive intestinal peptide, growth hormone releasing factor, melanocyte stimulating hormone, neurotensin, adrenal peptide, parathyroid hormone and related peptides, somatostatin and related peptides, brain natriuretic peptide, calcitonin, corticotropin releasing factor (CRF), cocaine amphetamine regulated transcript (CART), thymosin, urotensin, glucagon and glucagon like peptides (GLP-1 and GLP-2), somatostatin, interferon, a vascular endothelial growth factor (VEGF), platelet derived growth factor (PDGF), factor VII, factor VIII, factor V, factor IX, interleukins, urokinase, erythropoietin (EPO), chymosin, tissue plasminogen activator, CCK or serum albumin.

With specific regard to "glucagon and glucagon like peptides", this term as used herein may refer to polypeptides of human origin or from other animals and recombinant or semisynthetic sources and include all members of the glucagon family, such as GRPP (glicentine related polypeptide), glucagon, GLP-1 (glucagon like peptide 1), and GLP-2 (glucagon like peptide 2), including truncated and/or N-terminally extended forms, such as GLP-1(7-36), and includes analogues, such as GLP-1(7-35)R36A GLP-2 F22Y, GLP-2 A19T+34Y. GLP2 A2G and GLP-2 A19T, and other analogues having from 1 to 3 amino acid changes, additions and/or deletions.

Host Cells and the Expression of Recombinant Polypeptides Therefrom

The host cells for use in the present invention can be procaryotic or eucaryotic. The eucaryotic host cells for use in the present invention can be, for example, fungal, mammalian, plant or insect cells. Preferably, the host cells are yeast cells.

In order to produce the desired polypeptide, the host cell of the invention comprises a nucleic acid sequence encoding the recombinant polypeptide as well as regulatory sequences for directing the expression of the desired product such as regions comprising nucleotide sequences necessary or e.g. transcription, translation and termination. The genetic design of the host cell of the invention may be accomplished by the person skilled in the art, using standard recombinant DNA technology for the transformation or transfection of a host cell.

Preferably, the host cell is modified by methods known in the art for the introduction of an appropriate expression cassette in, for example a plasmid or a viral vector, comprising the nucleic acid encoding the recombinant polypeptide. The expression cassette may be introduced into the host cell by a number of techniques including, but not limited to, as an autonomously replicating plasmid or integrated into the chromosome.

Expression cassettes may contain regulatory sequences such as transcription control sequences, translation control sequences, origins of replication, and other regulatory sequences that are compatible with the recombinant cell and that control the expression of nucleic acid molecules encoding the recombinant polypeptide. In particular, recombinant nucleic acid molecules of the present invention include transcription control sequences. Transcription control sequences are sequences which control the initiation, elongation, and termination of transcription. Particularly important transcription control sequences are those which control transcription initiation, such as promoter, enhancer, operator and repressor sequences. Suitable transcription control sequences include any transcription control sequence that can function in at least one of the recombinant cells of the present invention. A variety of such transcription control sequences are known to those skilled in the art. Preferred transcription control sequences include those which function in bacterial, yeast, arthropod and mammalian cells, such as, but not limited to, tac, lac, trp, trc, oxy-pro, omp/lpp, rrnB, bacteriophage lambda, bacteriophage T7, T7lac, bacteriophage T3, bacteriophage SP6, bacteriophage SP01, metallothionein, alpha-mating factor, Pichia alcohol oxidase, alphavirus subgenomic promoters (such as Sindbis virus subgenomic promoters), antibiotic resistance gene, baculovirus, *Heliothis zea* insect virus, vaccinia virus, herpesvirus, raccoon poxvirus, other poxvirus, adenovirus, cytomegalovirus (such as intermediate early promoters), simian virus 40, retrovirus, actin, retroviral long terminal repeat, Rous sarcoma virus, heat shock, phosphate and nitrate transcription control sequences as well as other sequences capable of controlling gene expression in prokaryotic or eukaryotic cells. Transcription control sequences of the present invention are most preferably naturally occurring transcription control sequences associated with yeast. Suitable promoters for *S. cerevisiae* include the MFα1 promoter, galactose inducible promoters such as GAL1, GAL7 and GAL10 promoters, glycolytic enzyme promoters including TPI1 and PGK1 promoters, TRP1 promoter, CYC1 promoter, CUP1 promoter, PHO5 promoter, ADH1 promoter, and HSP promoter. A suitable promoter in the genus *Pichia* is the AOX1 (methanol utilisation) promoter.

Recombinant polypeptides of the present invention may also (a) contain secretory signals to enable an expressed polypeptide to be secreted from the cell that produces the polypeptide and/or (b) contain fusion sequences which lead to the expression of fusion proteins. Examples of suitable signal segments include any signal segment capable of directing the secretion of the fusion protein. Preferred signal segments include, but are not limited to, tissue plasminogen activator (t-PA), interferon, interleukin, growth hormone, histocompatibility and viral envelope glycoprotein signal segments, the leader sequence originating from the fungal amyloglucosidase (AG) gene such as galA - both 18 and 24 amino acid versions e.g. from *Aspergillus sp.*, the α-factor gene of yeasts e.g. from *Saccharomyces sp.* and *Kluyveromyces sp.*, the P-factor of *Schizosaccharomyces sp.*, and the α-amylase gene from *Bacillus sp.*, as well as natural signal sequences.

The cloning vehicle may also comprise a selectable marker, e.g. a gene, the product of which complements a defect in the host cell, or one which confers antibiotic resistance, such as ampicillin, kanamycin, chloramphenicol or tetracycline resistance.

The procedures used to ligate the DNA construct of the invention, the promoter, terminator and other elements, respectively, and to insert them into suitable cloning vehicles containing the information necessary for replication, are well known to persons skilled in the art.

Recombinant DNA technologies can be used to improve the expression of transformed nucleic acid molecules by manipulating, for example, the number of copies of the nucleic acid molecules within a host cell, the efficiency with which those nucleic acid molecules are transcribed, the efficiency with which the resultant transcripts are translated, and the efficiency of post-translational modifications. Recombinant techniques useful for increasing the expression of nucleic acid molecules useful for the methods of the present invention include, but are not limited to, operably linking the nucleic acid molecule to high-copy number plasmids, integration of the nucleic acid molecule into one or more host cell chromosomes, addition of vector stability sequences to plasmids, substitutions or modifications of transcription control signals (e.g., promoters, operators, enhancers), substitutions or modifications of translational control signals (e.g., ribosome binding sites, Shine-Dalgarno sequences), modification of nucleic acid molecule to correspond to the codon usage of the host cell, and the deletion of sequences that destabilize transcripts. The activity of an expressed recombinant polypeptide of the present invention may be improved by fragmenting, modifying, or derivatizing polynucleotide molecules encoding such a protein.

Methods of Producing Recombinant Polypeptides

Host cells that have been transfected or transformed with the nucleic acid encoding the recombinant polypeptide are cultured for a sufficient time and under appropriate conditions known to those skilled in the art in view of the teachings disclosed herein to permit the production, and preferably secretion, of the polypeptide, followed by recovery of the desired product.

Furthermore, owing to the reduced activity of the metalloprotease, the recombinant polypeptide expressed by the host cell may be obtained as a precursor protein, i.e. a zymogen, a hybrid protein, a protein obtained as a pro sequence or pre-pro sequence, or in unmaturated form.

The broth or medium used for culturing may be any conventional medium suitable for growing the host cell in question, and may be composed according to the principles of the prior art. The medium preferably contains carbon and nitrogen sources and other inorganic salts. Suitable media, e.g. minimal or complex media, are available from commercial suppliers, or may be prepared according to published protocols.

With regard to yeast host cells, it is often advantageous to produce heterologous polypeptides in a diploid yeast culture, because possible genetical defects may become phenotypically expressed in a haploid yeast culture, e.g. during continuous fermentation in production scale, and because the yield may be higher. The production of recombinant polypeptides in yeast host cell is described in Molecular Genetics of Yeast: A Practical Approach, Ed. J.R. Johnston, IRL Press (1994) which is incorporated herein by reference.

After cultivation, the protein is recovered by conventional methods for isolation and purification proteins from a culture broth. Well-known purification procedures include separating the cells from the medium by centrifugation or filtration, precipitating proteinaceous components of the medium by means of a salt such as ammonium sulphate, and chromatographic methods such as e.g. ion exchange chromatography, gel filtration chromatography, affinity chromatography, etc.

Examples

Materials and Methods

Yeast strains and growth conditions

The yeast strains used are listed in Table I. Construction of strains were carried out using either the two step gene disruption technique (Rothstein, 1991) or the PCR based method by (Brachmann et al., 1998). Media were purchased from Difco, amino acids and other supplements from Sigma-Aldrich. Yeast cells were grown at 30°C in YPD (1% yeast extract, 2% peptone and 2% dextrose) or synthetic complete media (SC) based on yeast nitrogen base with ammonium sulfate, succinic acid, NaOH and appropriate amino acids. Transformations with either linear DNA or plasmids were performed using the modified lithium acetate procedure as described (Gietz et al., 1995). Analysis of heterologous expressed CCK was performed from yeast growing in exponential phase due to the consistency in CCK-22 biosynthesis, in contrast to the results from yeast within the stationary phase (Fig. 2). ProCCK processing was analysed from cell extract and media of 5 A_{600} units of cells per ml synthetic complete media. Cell growth was followed by the absorbance at 600 nM.

Table 1. *S. cerevisiae* strains used in this study. Null mutants of putative metalloproteases are named by the ORF in the genotype and (*) represents mitochondrial proteases.

Strain	Genotype	Source
BY4705	*MATα ade2Δ· hisG his3Δ200 leu2Δ0 lys2Δ0 ura3Δ0*	(Brachmann et al., 1998)
LJY13	*MATα ade2Δ..hisG his3Δ200 leu2Δ0 lys2Δ0 ura3Δ0 yps1:.TRP1*	This study
LJY14	*MATα ade2Δ::hisG his3Δ200 leu2Δ0 lys2Δ0 ura3Δ0 yps1:.TRP1 yps3:··LEU2*	This study
LJY15	*MATα ade2Δ·:hisG his3Δ200 leu2Δ0 lys2Δ0 ura3Δ0 yps1::TRP1 yps3. LEU2 yps2··URA3*	This study
BJ2168	*MATa prc1-407 prb1-1122 pep4-3 leu2 trp1 ura3-52 gal2*	(Jones, 1991)
LJY21	*MATa prc1-407 prb1-1122 pep4-3 leu2 trp1 ura3-52 gal2 kex1. LEU2*	This study
LJY22	*MATa prc1-407 prb1-1122 pep4-3 leu2 trp1 ura3-52 gal2 kex1::LEU2 kex2::TRP1*	This study
LJY23	*MATa prc1-407 prb1-1122 pep4-3 leu2 trp1 ura3-52 gal2 kex2: TRP1*	This study
LJY122	*MATα ape1·:KANMX ape2·::LYS2 his3Δ0 leu2Δ0 lys2Δ0 ura3Δ0*	This study
LJY123	*MATα ape1::KANMX ape2::LYS2 ape3· LEU2 his3Δ0 leu2Δ0 lys2Δ0 ura3Δ0*	This study
LJY201	*MATa prc1-407 prb1-1122 pep4-3 leu2 trp1 ura3-52 gal2 axl1·:LEU2*	This study
LJY202	*MATa prc1-407 prb1-1122 pep4-3 leu2 trp1 ura3-52 gal2 ste24:.LEU2*	This study
LJY203	*MATa prc1-407 prb1-1122 pep4-3 leu2 trp1 ura3-52 gal2 prd1:LEU2*	This study
LJY204	*MATa prc1-407 prb1-1122 pep4-3 leu2 trp1 ura3-52 gal2 yil108w:.LEU2*	This study
Y15298	*MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 ste23··KANMX*	Euroscarf
Y11874	*MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 aap1: KANMX*	Euroscarf
Y10148 (*)	*MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 afg3. KANMX*	Euroscarf
Y14953	*MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 ape1.·KANMX*	Euroscarf
Y16224 (*)	*MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 rca1.·KANMX*	Euroscarf
Y14984 (*)	*MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 mip1::KANMX*	Euroscarf
Y17144	*MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 yme1::KANMX*	Euroscarf
Y13211	*MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 ybr074w:.KANMX*	Euroscarf
Y13801	*MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 ydl104c. KANMX*	Euroscarf
Y11941	*MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 yhr113w KANMX*	Euroscarf
Y11960	*MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 yhr132c.·KANMX*	Euroscarf
Y12296	*MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 yil137c··KANMX*	Euroscarf
Y15370	*MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 ynl045w:.KANMX*	Euroscarf
10864B	*MATα ura3-Δ851 leu2-Δ1 his3Δ200 lys2Δ202 ykr035c-ykr038c:.URA3*	Euroscarf
Y11749	*MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 yol057w::KANMX*	Euroscarf
Y16248	*MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 yol098c::KANMX*	Euroscarf
10231B	*MATα his3Δ200 leu2Δ1 trp1Δ63 ura3-52 yol154w(4,744):.KANMX*	Euroscarf
Y14266	*MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 ydr430c::KANMX*	Euroscarf
LJY430	*MATa prc1-407 prb1-1122 pep4-3 leu2 trp1 ura3-52 gal2 ydr430c..LEU2*	This study
LJY432	*MATa prc1-407 prb1-1122 pep4-3 leu2 trp1 ura3-52 gal2 kex2:·TRP1 ydr430c··LEU2*	This study

DNA extraction and amplification

Yeast genomic DNA was isolated as described (Philippsen et al., 1991). Polymerase chain reaction (PCR) was performed using either *Pwo* polymerase or the enzyme cocktail based on *Taq*, *Pwo* and *Pfu* polymerase (Expand long range PCR kit, XL-PCR) both from Roche. All PCR products were visualised by agarose gel-electrophoresis and PCR products either purified from the gel using the gel-extraction kit (Qiagen) or from the reaction mixture by PCR purification spin columns (GENOMED). PCR based one step gene disruption was performed using 50 ng of plasmid from the pRS400 series (Brachmann et al., 1998) as template. Amplification of the marker was performed with oligonucleotides having 20 nucleotides towards the plasmid and additional 50 nucleotides flanking the target gene (Table 2). All other DNA manipulations were carried out according to standard procedures (Sambrook et al., 1989).

Plasmid constructions

Expression of proCCK was performed in pRS426 [2µ *URA3*] (Brachmann et al., 1998) using the phosphoglycerate kinase promoter (*PGK1p*) and terminator (*PGK1t*). The *PGK1* promoter was amplified with *PGK1p* 5'*Hind*III and *PGK1p* 3'*MCS* (Table 2) using 100 ng of genomic yeast DNA as template and subsequently cloned into pGEM-11 (Promega) in the *Hind*III and *Sac*I restriction enzyme sites. The terminator was amplified with *PGK1t* 5'*Bgl*II and *PGK1t* 3'*Sac*I (Table 2) and ligated into the plasmid containing the promoter at the *SacI* and *EcoR*I restriction enzyme sites. This construct, pGEM-11 *PGK1pMCSPGK1t* then contained the *PGK1*-promoter, a multiple cloning site (*MCS*) with the restriction enzyme sites *EcoR*I, *BamH*I, *Xba*I and *Bgl*II followed by the *PGK1* terminator. The preproMfα1p-proCCK fusion (Rourke et al., 1997) (Fig. 1) was subcloned into the *EcoR*I and *Xba*I sites of pGEM-11 *PGK1pMCSPGK1t* and finally the entire gene was cloned into pRS423 as well as pRS426 to complete the yeast CCK expression constructs, pRS423 preproMfα1p-proCCK and pRS426 preproMfα1p-proCCK respectively. Expression of *CYM1* on a multi copy plasmid was constructed by amplification of the open reading frame (ORF) of *CYM1* and additional 926 bp at the 5' end and 703 bp at the 3' end. The amplification was carried out by XL-PCR using 100 ng of genomic yeast DNA and the oligonucleotides, *CYM1* 5'*Apa*I and

Table 2. Oligonucleotides used.

Oligo	Oligonucleotide sequence (5'-3')	Purpose
PGK1p5'HindIII	AATAGAAGCTTGTCGACTGATCTATCCAAAACTG (SEQ ID NO: 16)	Expression construct
PGK1p3'MCS	AAAAGAGCTCGGCCAGATCTTCTAGAGGATCCAA GAATTCTGTTTTATATTTGTTGTAAAAAGTAG (SEQ ID NO: 17)	Expression construct
PGK1t5'BglII	TTTTGAATTCCAAGATCTCCCATGTCTCTACTGGTGG (SEQ ID NO: 18)	Expression construct
PGK1t3'SacI	CCCCGAGCTCGTCGACCCTTCTCGAAAGCTTTAACGAACGC (SEQ ID NO: 19)	Expression construct
5'MFα1-*EcoRI*	TTTTGAATTCAAAGAATGAGATTTCCTTCAATTTTTACTGCAG (SEQ ID NO: 20)	preproMfα1p-proCCK
CCK3'-*XbaI*	TTTTTCTAGACTAGGAGGGGTACTCATACTCCTCGGC (SEQ ID NO: 21)	preproMfα1p-proCCK
CCK-22 K→A (S)	CGAATGTCCATCGTTGCGAACCTGCAGAACCTG (SEQ ID NO: 22)	Lys^{61}→Ala^{61} mutation
CCK-22 K→A (AS)	CAGGTTCTGCAGGTTCCTAACGATGGACATTCG (SEQ ID NO: 23)	Lys^{61}→Ala^{61} mutation
CCK-22 K→R (S)	CGAATGTCCATCGTTAGGAACCTGCAGAACCTG (SEQ ID NO: 24)	Lys^{61}→Arg^{61} mutation
CCK-22 K→R (AS)	CAGGTTCTGCAGGTTCCTAACGATGGACATTCG (SEQ ID NO: 25)	Lys^{61}→Arg^{61} mutation
CCK-22 seq	TCGCAGAGAACGGATGGC (SEQ ID NO: 26)	Sequencing
CYM15'ApaI	TTTTGGGCCCTTCATGGTGATACGGTATCTCTTGGC (SEQ ID NO: 27)	Cloning of *CYM1*
CYM13'XhoI	TTTTCTCGAGAAGGTGGAACATACTGCCCTGGGATGG (SEQ ID NO: 28)	Cloning of *CYM1*
KEX25'	TTTTGAGCTCGTTTAGGAAACGTCCTTGGCGGAGATGC (SEQ ID NO: 29)	Cloning of *KEX2*
KEX23'	TTTTTCTAGACACTGCGAATCCATGGTATAAACCAAAACC (SEQ ID NO: 30)	Cloning of *KEX2*
*KEX2*DC5'	GTCGTTGTTCATGGACATACCTCC (SEQ ID NO: 31)	Control of Δ*kex2*
*KEX2*DC3'	TACAAATGTTCTTCTGCCATTTCTGG (SEQ ID NO: 32)	Control of Δ*kex2*
TRP15'NdeI	GGTTCATATGCGCCGGAGCTCCTCGACAGCAG (SEQ ID NO: 33)	Cloning of *TRP1*
TRP13'AvrII	GGTTCCTAGGATCCGCAAGTTTGATTCCATTGCGGTG (SEQ ID NO: 34)	Cloning of *TRP1*
KEX15'GD400	TTAAAGAGTACCTTGGCTATAGAATACCGTAGAGATAAAGA CCTGAATAGAGATTGTACTGAGAGTGCAC (SEQ ID NO: 35)	*KEX1* deletion
KEX13'GD400	AGGTATTATAACTATTTTTCTGTATTTTTTATATATTTTTATTT GCCAAGCTGTGCGGTATTTCACACCG (SEQ ID NO: 36)	*KEX1* deletion
KEX15'DC400	CTTTGGTTAAAGAGTACCTTGGC (SEQ ID NO: 37)	Control of Δ*kex1*
KEX13'DC400	TACTACGAAAAGCGTGTGCGAGG (SEQ ID NO: 38)	Control of Δ*kex1*
CYM15'GD400	TAGAAGGCTACTCAAAAGAATAAAGTTACTATAAAATATACT GCGGTATATAGATTGTACTGAGAGTGCAC (SEQ ID NO: 39)	*CYM1* deletion

*CYM1*3'GD400	GATCGGCAAGAAACTTTGAAGCAGTATATTTACAGGATTAAA TTATATATCTGTGCGGTATTTCACACCG (SEQ ID NO: 40)	*CYM1* deletion
*CYM1*5'DC400	CGGAGGGGCTCTATGATAAAGG (SEQ ID NO: 41)	Control of Δ*cym1*
*CYM1*3'DC400	GAGTAACTAGGGCTTCTCTTCCC (SEQ ID NO: 42)	Control of Δ*cym1*

CYM1 3'*Xho*I (Table 2). The PCR product was purified on spin columns and subsequently cloned into the *Apa*I and *Xho*I restriction enzyme sites of pRS425.

The Lys^{61} residue, believed to be crucial for the proteolysis of proCCK to release CCK-22, was exchanged by Ala by site-directed mutagenesis (Horton et al., 1993). The exchange was performed by PCR using the *Pwo* polymerase (Boehringer Mannheim), where two products were amplified with the oligonucleotides sets, *PGK1p*5' *Hind*III / CCK-22 K→A (antisense) and CCK-22 K→A (sense) / *PGK1t*3' *Sac*I (Table 2) and 50 ng of pRS426 preproMfα1p-proCCK as template to each reaction. The two products were subjected to agarose gel-electrophoresis and approximately 1 mm^2 of each product where cut out and used directly as template in a third PCR reaction. In this reaction the full-length cDNA encoding the fusion protein was amplified using *PGK1p*5' *Hind*III and *PGK1t*3' *Sac*I (Table 2). The PCR product was subcloned into pCR-Blunt II (Invitrogen) and sequenced with the CCK specific primer, CCK-22 seq. Finally the *PGK1p* preproMfα1p-proCCK (K→A) *PGK1t* product was cloned into the *Hind*III and *Sac*I sites of pRS426 to construct the expression plasmid, proCCK (K→A). Substitution of Lys with Arg was performed as described above by exchanging the CCK specific primers with CCK-22 K→R (antisense) and CCK-22 K→R (sense) to construct the proCCK (K→R) vector.

Strain construction

Construction of a partial *KEX2* disruption was performed in BJ2168 by amplification of the entire *KEX2* gene with 1000 bp on each site of the ORF by XL-PCR using 100 ng of genomic yeast DNA and the oligonucleotides, *KEX2* 5'*Sac*I and *KEX2* 3'*Xba*I (Table 2). The PCR product was purified on spin columns and cloned into pCR-Blunt II (Invitrogen). Amplification of *TRP1* was performed by XL-PCR introducing an *Nde*I site 925 bp 5' to the ORF and an *Avr*II site 212 bp 3' to the stop codon using *TRP1* 5'*Nde*I and *TRP1* 3'*Avr*II (Table 2). The PCR product was purified and subcloned into the *Nde*I and *Avr*II sites of *KEX2* eliminating 2018 bp of *KEX2* and 170 bp of the promoter. The *kex2::TRP1* construct was excised from pCR-Blunt II using the *Not*I and *Spe*I restriction

enzymes and subsequently transformed into BJ2168. Transformants were selected on SC-Trp plates followed by a colony PCR screen to test for correct integration using oligonucleotides that cover the entire marker plus an additional 1200 bp on each site of *KEX2*, *kex2* DC5' and *kex2* DC3' (Table 2). Construction of a *kex2 kex1* strain was performed by the two step gene disruption technique (Rothstein, 1991) using the *LEU2* marker. Amplification of *LEU2* was performed by XL-PCR using 50 ng of pRS405 as template and *kex1*5'GD400 and *kex1*3'GD400 (Table 2). The PCR product was purified using PCR purification spin kit (GENOMED) and subsequently transformed into LJY23. Transformants were selected on SC-Leu plates and correct integration was tested by PCR-based colony screen using *kex1*5'DC and *kex1*3'DC (Table 2).

A Δ*cym1::LEU2* (LJY430) strain in a BJ2168 background was constructed by the one step gene disruption technique as described above for the Δ*kex1* strain using the oligonucleotides, *CYM1*5'GD400 and *CYM1*3'GD400 (Table 2) for gene disruption and *cym1*5'DC and *cym1*3'DC (Table 2) for disruption control. All null mutants created by this method were prepared with oligonucleotides designed towards the 50 bases adjacent to the 5' and 3' UTR with a specific 3' end to the pRS400 series of vectors containing various markers (Brachmann et al., 1998). Transformants were selected on appropriate agar plates followed by a colony PCR screen to test for correct integration using oligonucleotides that cover the entire marker plus an additional 200 bp on each site. Only the oligonucleotides that are not positioned as described above are shown in Table 2.

Gene deletions of *STE24*, *AXL1*, *PRD1* and *YIL108w* were made in BJ2168 using the PCR disruption technique (Brachmann et al., 1998) and pRS405 [*LEU2*] as template.

The LJY123, which contain gene deletions of *APE1*, *-2* and *-3*, was derived from Y14953 using PCR disruption technique (Brachmann et al., 1998). *APE2* was initially replaced with the *LYS2* (pRS317 [*cen; LYS2*]) where the PCR product was purified from agarose gel prior to transformation and *APE3* was substituted with the *LEU2* marker (pRS405 [*LEU2*]).

The *yps1 yps2 yps3* triple mutant (LJY15) was constructed in BY4705 using the PCR disruption technique (Brachmann et al., 1998). The ORF of *YPS1* were initially deleted by insertion of the *TRP1* locus (pRS404) to generate LJY13. This strain was then used as host for the deletion of *YPS3* by insertion of the *LEU2* marker (pRS405) and finally the *YPS2* was deleted by insertion of the *URA3* marker by amplification of pRS406 [*URA3*] to construct LJY15 (Table I).

CCK and CYM1 expression

Human proCCK was expressed as a fusion protein between the prepro leader sequence of yeast α-mating factor and proCCK (preproMfα1p-proCCK). The fusion construct was expressed on multi-copy plasmids, with constitutively gene transcription from the phosphoglycerate kinase promoter. "ProCCK expression" refers to expression using pRS426 preproMfα1p-proCCK, which was used in all yeast strains with exception of BY4705 and the isogenic yapsins deletion strains, where proCCK was expressed from pRS423 preproMfα1p-proCCK. *CYM1* expression was driven by its own promoter. Plasmid constructs, and oligonucleotides used are listed in Table 2.

Enzymatic treatment

Trypsin treatment was performed using 1 mg/ml Trypsin (Worthington Biochemical Corporation) in a 50 mM sodium phosphate buffer (pH 7.5) for 30 min at RT and terminated by immersion into boiling water for 10 min. Carboxypeptidase B (Boehringer Mannheim) treatment with a final concentration of 4 µg/ml was performed in 0.1 mM sodium phosphate buffer (pH 7.5) at room temperature for 30 min. The reaction was terminated by immersion into boiling water for 10 min.

Gel chromatography

Yeast transformants grown to late exponential phase were centrifuged at 15000 *g* to collect the cells and 500 µl of the medium was loaded directly onto a Sephadex G-50 superfine (Pharmacia) column (1×100 cm) at 4°C. The sample was eluted in VBA buffer (20 mM barbital buffer, 0.11% bovine serum albumin and 0.6 mM thiomersal) at a flow rate of 3.5 ml/h and fractions were collected every 17 min. Calibrations were performed by including ^{125}I-albumin (V_0) and ^{22}NaCl (V_t). The elution constants K_d, of peaks eluting at V_e are calculated as $K_d = (V_e-V_0)/(V_t-V_0)$.

Radioimmunoassay

Two different antisera were used to determine the amount of processed cholecystokinin. Ab 89009 (Paloheimo et al., 1994) is specific for the N-terminus of CCK-22 and Ab 7270 (Hilsted et al., 1986) is specific for Gly-extended CCK. The fraction of CCK processed to CCK-22 is calculated by division of the immunoreactivity measured with Ab 89009 with the amount measured with the same antibody after the

sample was treated with trypsin to measure the total amount of N-terminal extended CCK-22.

Yeast extract and protease assay

Ten A_{600} units of yeast cells growing in exponential phase were sedimented by centrifugation at 3000 *g* for 5 min, washed once in 25 ml H_2O and transferred to a 2 ml Eppendorf tube. An equal amount of acid washed glass beads (Sigma-Aldrich) was added followed by 200 µl of 0.1 M NaH_2PO_4 (pH 4.5) including various inhibitors (150 µM Bestatin, 30 µM E-64, 10 µM Leupeptin, 1 µM Pepstatin A, 1 mM phenylmethylsulfonyl fluoride (PMSF), 1 mM EDTA and 1 mM 1,10-orthophenanthroline or 1 tablet complete inhibitor with or without EDTA per 2.5 ml 0.1 M NaH_2PO_4 (Boehringer Mannheim). The cells were broken by vortexing 3×20 sec and the extracts were clarified by centrifugation at 15000 *g* for 10 min. All steps were carried out at 4°C. The protease assay was performed using 20 pmol synthetic amidated CCK-33 (Peninsula Laboratorie Europe, Merseyside, England) or Ac-CCK-33-Gly (Cambridge Research Biochemicals, Stockton, England) as substrate, 20 µl yeast extract, various inhibitors and activators in a total volume of 30 µl. The mixture was incubated at 30°C for 1 h and the reaction terminated by adding 500 µl VBA buffer followed by immediate immersion into a boiling water bath for 10 min.

Protease assay using metalloprotease deficient strains

The assay was performed as described above, but with addition of 1 mM Bestatin and 1 mM Mn^{2+} to decrease N-terminal degradation.

Protease assay using intact yeast cells

Five A_{600} units of exponential growing cells where sedimented, washed once in 5 ml H_2O and once in SC media (pH 6.0), before the cells where resuspended in 25 µl SC media. The protease assay was performed by addition of 20 pmol synthetic Ac-CCK-33-Gly as substrate and the mixture incubated with gentle shaking at 30°C for 1 h. The reaction was terminated by addition of 500 µl VBA and the cells removed by centrifugation before the supernatant was immersed into boiling water for 10 min.

Analysis of secreted CCK by MALDI-TOF

Fifty A_{600} units of CCK transformed yeast were subjected to 25 ml of fresh media, followed by inoculation for 3 h. Cells were removed by centrifugation at 15000 *g* for 10 min and 500 µl of media was concentrated and desalted by reverse phase using a ZipTip C_{18} column (Millipore). The peptides were eluted with 10 µl 50% acetonitrile. The purified peptides were analysed in a Matrix Assisted Laser Desorption/Ionization time-of-flight mass spectrometer (Biflex, Bruker-Franzen, Bremen, Germany) operated in the reflected mode using time lag focusing (delayed extraction). For analysis, 0.5 µl of the sample was mixed with 0.5 µl matrix solution (α-cyano-4-hydroxycinnamic acid in acetonitrile/methanol, Hewlett Packard). Then 0.5 µl of the mixture was applied to the probe and allowed to dry before introduction into the mass spectrometer.

Statistical analysis

Statistical calculations were performed using an unpaired students t-test to analyse whether the change in proCCK expression or the fraction of mature CCK-22 between wild type yeast expressing proCCK and mutants isogenic to the wild type strain can be considered to be statistically significant. The P-value calculated for CCK-22 processing between yapsin mutants are comparisons of BY4705 and each mutant, whereas the brackets represents comparisons between BY4705 *yps1* and BY4705 *yps1 yps3* and, *BY4705 yps1 yps3* and BY4705 *yps1 yps2 yps3*, respectively.

Results

The influence of growth conditions on the CCK-22 processing

The intra- and extracellular fraction of CCK-22 was measured from BJ2168 expressing proCCK. The intracellular fraction remained unaltered whether the cells were in exponential growth or had reached stationary phase (Fig. 2). However, the relative amount of secreted CCK-22 changed dramatically when the cells reached stationary phase. During exponential growth the fraction of CCK-22 was 23%, but in the stationary phase (after 270 min) the fraction increased to 37% (Fig. 2). Hence, for the experiments described herein only exponentially growing cells were used.

The significance of the Lys residue in the release of CCK-22

To evaluate the role of the Lys residue in proteolysis, transformants of BJ2168 with the two expression constructs, proCCK and proCCK (K→A) were grown to late exponential phase and the culture media collected. The media from each strain was subjected to gel chromatography and the content of Gly-extended CCK in the collected fractions where measured with Ab 7270. CCK from the wild type media eluted in two major peaks at K_d = 0.8 and 1.1 (Fig. 3 A) in accordance with the previously established elution positions for CCK-22-Gly and CCK-8-Gly (Cantor et al., 1987; Rourke et al., 1997), while the proCCK (K→A) only gave rise to CCK-8-Gly and a larger form eluting at a K_d = 0.6 (Fig. 3 C). The two peaks eluting at K_d = 0.7 and 0.8 for the wt construct (Fig. 3 B) correspond to C-terminally extended CCK-22 and CCK-22-Gly, respectively (Rourke et al., 1997), whereas no CCK-22 immunoreactivity was observed in these positions for the proCCK (K→A) construct (Fig. 3 D). However, a small peak of immunoreactivity was seen at K_d = 0.55 which may be due to the slight cross reactivity of Ab 89009 with a larger unprocessed fragment (Paloheimo et al., 1994). To investigate the effect of substituting Arg for Lys, proCCK (K→R) was transformed into BJ2168. Media from transformants where analysed before and after tryptic cleavage. The fraction of proCCK processed to CCK-22 was similar to that seen for wild type CCK (Fig. 11).

Analysis of secreted CCK peptides by mass spectrometry

Media collected from BJ2168 transformed with proCCK were analysed by mass spectrometry (Figure 12). The fragments obtained correspond to the processing leading to CCK-39, CCK-22 and CCK-8. Two peptides were identified N-terminal of Lys^{61} (Tyr^{45}-Val^{60}, 1805.0 Da and Tyr^{45}- Lys^{61}, 1932.2 Da). It appeared likely that the former was a carboxypeptidase degradation product of the latter. To elucidate this question and in an attempt to identify the C-terminal extended CCKs, the present inventors produced a disruption strain in which both *KEX2*, encoding the serine protease responsible for the processing to CCK-8, and the carboxypeptidase encoded by *KEX1* were mutated. Following transformation of proCCK into this *kex2 kex1* strain (LJY22) and subsequent analysis of the secreted peptides the inventors found only the peak corresponding to Tyr^{45}-Lys^{61}. The same pattern, with only the peak corresponding to Tyr^{45}-Lys^{61} was seen using single gene disruption of *KEX1* and *KEX2* to express proCCK (data not shown). Thus the Tyr^{45}-Val^{60} must be a degradation product in accordance with CCK-22 arising from cleavage after Lys^{61}. Additional fragments were discovered by CCK expression in the *kex2 kex1* strain corresponding to processing leading to CCK-61 (not identified in mammals), CCK-58, C-terminal extended CCK-39 and C-terminal extended CCK-22 (Figure 12), whereas none of the peptides corresponding to CCK-8 could be

identified, in accordance with our previous work showing that Kex2p is responsible for this processing (Rourke et al., 1997).

Kex2p is involved in the biosynthesis of CCK-22

Previous analysis of CCK peptides secreted from a *kex2* strain as well as the results obtained by mass spectrometry indicate that the cleavage at Lys^{61} releasing CCK-22 can occur without the involvement of Kex2p. However, the *kex2* strain, shows a decrease in CCK-22 concentration. ProCCK was expressed both in the vacuole protease deficient and the isogenic *kex2* strain (BJ2168 and LJY23) and the processed intra- and extracellular fractions of CCK-22 from exponentially growing cells were measured. Approximately 28% of the intracellular CCK content was processed after Lys^{61} in BJ2168, whereas only 6% was processed within the *kex2* strain. Analysis of secreted CCK peptides showed that the media collected from BJ2168 expressing proCCK contained approximately 20% CCK-22, whereas from the *kex2* mutant, the amount was reduced to 5%. These results indicate that Kex2p is involved in the processing leading to CCK-22. However, there are other proteases that can perform the cleavage at Lys^{61}.

In vitro assay of Lys^{61} cleavage

To investigate the nature of the protease(s) in addition to Kex2p that are able to perform the endoproteolytic cleavage after the single Lys^{61} residue of proCCK, an *in vitro* assay was established using crude preparations from of *S. cerevisiae* and synthetic CCK-33 as substrate.

Using extract from the vacuole protease deficient strain, BJ2168, there was an extensive N-terminal degradation, and the recovery of measurable CCK was less than 10% of the control without yeast extract. Because the assay depends on the intact N-terminus of CCK-22 for the antibody to bind, the inventors created a strain where some of the known *S. cerevisiae* aminopeptidases were deleted. The Y14953 strain (*ape1*) was used as parental strain in which the *APE2* and *APE3* genes were also deleted. Using this LJY123 strain to prepare the cell extract there was a 2-3 fold better recovery of immunoreactivity compared to the recovery seen with BJ2168.

Processing to CCK-22 depends on metal ions

The nature of the protease performing the cleavage of synthetic human CCK-33 to CCK-22 was analysed by inclusion of a number of different inhibitors with the extract from LJY123. The results showed that only the addition of a metal chelating agent inhibited proteolysis of CCK-33 to CCK-22 (Fig. 4 A).

The metal dependency of the protease was tested *in vitro*, after the activity initially was inhibited by addition of 1 mM EDTA. Reconstitution of the activity leading to maturation of CCK-22 was tested by addition of different divalent cations in 0.2 mM surplus. Addition of Zn^{2+}, Co^{2+} and Mn^{2+} could reestablish the protease activity, whereas Ca^{2+}, Cu^{2+} or Mg^{2+} had no effect (Fig. 4 B) in accordance with the properties of known metalloproteases, which are only activated by Zn^{2+}, Co^{2+} and Mn^{2+}. Reactivation using increasing Zn^{2+} concentrations showed a biphasic pattern, with Zn^{2+} acting inhibitory at concentrations above 5 mM (data not shown).

The time course of CCK-cleavage by Zn^{2+} and Mn^{2+} reactivated metalloproteases were analysed using cell extract from LJY123, after initial inhibition with 1 mM EDTA. Reactivation was performed by addition of 1.2 mM Zn^{2+} or Mn^{2+} followed by incubation for 30, 60 and 120 min. In this assay and the following *in vitro* protease assays the inventors used the N-terminal acetylated CCK-33-Gly (Ac-CCK-33-Gly) as substrate, which resulted in much slower non-specific degradation. Measurement of the CCK-22 immunoreactivity before and after tryptic cleavage using Ab 89009 showed no difference in the activation potency between Zn^{2+} and Mn^{2+} at 30 and 60 min, however after 120 min 10% more CCK-22 immunoreactivity was measured using Mn^{2+} as activator compared to Zn^{2+} (Fig. 5). This increase in immunoreactivity is probably due to an inhibition of degradation following addition of Mn^{2+} here as well as to the yeast cell extracts used in Table 3.

Table 3. Metalloproteases in *Saccharomyces cerevisiae*. Search performed in Swiss-Prot Sequence Retrieval System (SRS) http://www.expasy.ch/. Protease assay performed in two independent assays (A and B) using extracts from the metalloprotease deficient strains. The amount of CCK-22 is measured with Ab 89009 and the total amount of CCK is measured after tryptic cleavage with Ab 89009. Putative metalloproteases are marked with *.

Name	Swiss-Prot acc #	ORF	CCK-22 [nM]		Total CCK [nM]		Fraction CCK-22	
			A_1	B_1	A_2	B_2	A_1/A_2	B_1/B_2
AAP1	P37898	YHR047c	3.2	3.2	36	32	0.09	0.10
AFG3	P39925	YER017c	2.8	2.4	23	24	0.12	0.10
APE1	P14904	YKL103c	4.4	3.7	34	28	0.13	0.13
APE2	P32454	YKL157w						
APE3	P37302	YBR286w						
DPP3	Q08225	YOL057w	4.3	4.6	31	35	0.14	0.13
LTA4	Q10740	YNL045w	3.9	4.0	32	29	0.12	0.13
MIP1	P35999	YKL134c	3.4	3.3	34	33	0.10	0.10
PRD1	P25375	YCL057w	2.4	2.8	28	26	0.09	0.11
QRI7*	P43122	YDL104c	2.8	2.9	23	24	0.12	0.12
RCA1	P40341	YMR089c	4.2	3.5	31	27	0.14	0.13
STE23	Q06010	YLR389c	2.6	2.3	25	20	0.10	0.12
STE24	P47154	YJR117w	3.4	2.8	19	17	0.18	0.16
YBS4*	P38244	YBR074w	2.6	2.7	27	29	0.10	0.09
YHR3*	P38821	YHR113w	2.5	2.4	26	26	0.10	0.09
YHT2*	P38836	YHR132c	3.8	3.3	34	32	0.11	0.10
YIK8*	P40483	YIL108w	5.7	5.2	43	39	0.13	0.13
YIN7*	P40462	YIL137c	3.5	2.8	23	20	0.15	0.14
YK18*	P36132	YKR038c	2.9	2.1	23	20	0.13	0.11
YME1	P32795	YPR024w	2.3	2.7	25	25	0.09	0.11
MAS2	P11914	YHR024c	ND, Lethal genes					
MAS1	P10507	YLR163c						
AXL1	P40851	YPR122w	3.5	3.4	31	30	0.11	0.11
CYM1*	P32898	YDR430c	0.6	0.4	42	39	0.01	0.01
YOJ8*	Q12496	YOL098c	3.6	3.8	35	34	0.10	0.12

Extracellular yapsin activity

To investigate whether any protease activity is secreted or attached extracellular to the plasma membrane, the protease activity was assayed in media and with intact yeast cells. No degradation of CCK-33 occurred after 1 h of incubation at 30°C using media from exponential growing LJY123 cells in accordance with earlier observations (Rourke et al., 1997). During incubation with intact yeast cells, cleavage to expose the N-terminus of CCK-22 could be measured (Fig. 6) however, this protease activity could not be abolished by the inhibitors investigated (data not shown). By using intact cells containing gene disruptions of *YPS1, YPS2* and *YPS3* (Table I) the fraction of processed CCK-22 decreases by deletion of each of the three aspartyl proteases compared to wild type cells (Fig. 6). These data show that the three proteases all have extracellular protease activity, which can cleave at Lys^{61} in proCCK. Preliminary results indicate that gene disruption of *YPS7* decreases extracellular Lys^{61} processing in amounts comparable to the *YPS1* deletion (unpublished results).

Expression of proCCK in metalloprotease deficient strains

Based on previously described metalloproteases in *S. cerevisiae* with endoproteolytic activity (Adames et al., 1995; Schmidt et al., 2000), gene deletion strains of *AXL1* (LJY201) and *STE24* (LJY202) were initially prepared in BJ2168. ProCCK expression in these strains showed that proteolysis after Lys^{61} was unchanged compared to wild type, and it was decided to test the remaining metalloprotease deficient strains LJY123, LJY203, LJY204 and the metalloprotease deficient strains obtained through Euroscarf (Table I) for their ability to secrete CCK-22 (mitochondrial peptidases were not included). The CCK-22 immunoreactivity did not change significantly among the CCK producing metalloprotease deficient strains (data not shown), and no protease responsible for the processing of heterologous expressed proCCK to CCK-22 was identified by this approach.

CYM1 encodes a protease that can release the free N-terminus of CCK-22

Cell extracts were prepared from each of the viable metalloprotease deficient strains and tested in the *in vitro* protease assay to investigate whether any reduction in proteolysis was measurable. In this assay 1 mM Mn^{2+} and 1 mM bestatin were included prior to the addition of Ac-CCK-33-Gly, since it was found that the recovery was 80-90% compared to 30% without addition of these aminopeptidase inhibitors (data not

shown). Deletion of *CYM1* almost abolished the protease activity, whereas none of the other metalloprotease deficient strains showed a significant change in the biosynthesis of CCK-22 (Table 3).

Expression of *CYM1* on a multicopy plasmid increase the fraction of matured CCK-22 in vitro

To determine whether the amount of synthetized CCK-22 correlates with the amount of Cym1p *in vitro*, Cym1p was expressed on a multicopy plasmid and the fraction of synthetized CCK-22 analysed over time. Cell extract from BJ2168 transformed with pRS425 *CYM1* and the control transformed with the empty pRS425 vector were used in the *in vitro* protease assay with 1 mM Mn^{2+} in which the reactions were terminated after 15, 30, 45 and 60 min. The CCK-22 immunoreactivity was measured with Ab 89009 and the remaining CCK-33 was measured with the same antibody after tryptic cleavage (Fig. 7). Expression of *CYM1* on a multicopy plasmid enhanced the rate of CCK-22 production several fold. However, the inventors also observed an increased degradation of CCK-33 and CCK-22 (Fig. 7 B). When the same experiment was performed at pH 6.0 and pH 7.5, there was a dramatically increased degradation and after 30 min incubation the CCK immunoreactivity was undetectable at pH 6.0 (data not shown). These results show the Lys-specific cleavage in CCK-22 maturation *in vitro* is dependent on the amount of Cym1p.

Expression of proCCK in *cym1* mutant strain enhance CCK secretion

To elucidate the role of *CYM1* in the biosynthesis of CCK-22 *in vivo*, gene deletions of *CYM1* were prepared in the vacuole protease deficient strain, BJ2168, and isogenic *kex2* strain. Deletion of *CYM1* resulted in an approximately 40% increase in the total amount of proCCK within the cells (Fig. 8 A) accompanied by a similar decrease in CCK-22 independent of *KEX2* disruption (Fig. 8 C). Also the secreted amount of total CCK in the *cym1* strains increased with more than 60% (Fig. 8 B), but unlike the fractional decrease in intracellular CCK-22 there was an increase in the extracellular fractions of CCK-22 compared to vacuole protease deficient strain and the isogenic *kex2* strain (Fig. 8 D).

Expression of CCK K→A mutant leads to intracellular CCK accumulation comparable to the accumulation of wild type CCK in a cym1 strain

The observations that a gene disruption of *CYM1* causes an increase in intracellular concentrations of CCK (Fig. 8 A) raises the question whether the proteolytic activity of Cym1p leads to degradation of CCK-22 prior to translocation into the ER. Therefore, the inventors examined the intracellular CCK content in strains expressing CCK where the maturation of CCK-22 has been eliminated by using the $Lys^{61} \rightarrow Ala^{61}$ mutant. Transformants of this CCK mutant in the vacuole protease deficient strain, BJ2168 and the isogenic *cym1* strain were analysed using Ab 7270 after trypsin and carboxypeptidase B treatment and there was an increase in the intracellular CCK immunoreactivity for this construct compared to expression of wild type CCK (Fig. 9). Mutant CCK (K→A) and wild type CCK transformants resulted in an increase in the intracellular proCCK concentration when expressed in BJ2168 and the *CYM1* disruption strain, respectively. The increase in intracellular proCCK was not additive showing that proteolytic activity of Cym1p leads to degradation of CCK-22 prior to translocation into the ER.

Expression of proCCK in aspartyl protease deficient strains

The Lys^{61}-specific cleavage of proCCK was analysed in null mutants of *YPS1*, *-2* and *-3*, where the intra- and extracellular amount of CCK-22 was measured from exponentially growing cells of wild type yeast, BY4705 and the isogenic aspartyl protease deficient strains transformed LJY13, -14 and -15 with proCCK. Both intra- and extracellular CCK immunoreactivity of BY4705 was lowered more than 10 fold compared to the vacuolar protease deficient strain, BJ2168 (data not shown). The intracellular fraction of CCK-22 decreased significantly from approximately 28% in wild type cells to 17% in the *yps1* strain, whereas no additional decrease could be measured by gene disruptions of *YPS2* and *YPS3* (Fig. 10 A). The extracellular fraction of CCK-22 did however show that Yps1p, Yps2p and Yps3p all are involved in the biosynthesis of CCK-22 and that the triple mutant reduced the fraction of CCK-22 to 2/3 compared to wild type yeast (Fig. 10 B).

CYM1 disruption leads to a two-fold increase in the total amount of secreted wild type CCK as well as the CCK K→R mutant

To elucidate whether Cym1p cleaves C-terminally to a single Arg residue, the CCK (K→R) mutant was expressed in the vacuolar protease deficient strain, BJ2168 and the isogenic *cym1* strain. The concentration of both intra- and extracellular CCK was compared to wild type CCK expressed in these strains. The total amount of the mutated CCK (K→R) was increased both intra- and extracellular comparable to wild type CCK

(Fig. 11). Both wild type CCK and the $Lys^{61} \rightarrow Arg^{61}$ mutant showed more than a two-fold increase in the measurable amount of extracellular CCK when expressed in the *cym1* strain (Fig. 11 B).

Usage of Cym1p activity in the synthesis of peptides

Another aspect of the invention is to use the activity from Cym1p, either expressed from its own promoter or from a strong constitutive promoter such as *PGK1*, *ADH1* or *TPI1*, or the induceable *GAL1* promoter to produce an increased amount cytosolic Cym1p activity. As previously mentioned, the synthesis of CCK-22 is significantly increased when *CYM1* is transcribed from its own promoter on a 2μ plasmid (Fig. 7). Transcription can either be performed from a plasmid containing the promoter, *CYM1* and a terminator, or by introducing the desired promoter into the genome by heterologous recombination to substitute the endogenous promoter of *CYM1*.

The activity can be used intracellulary to generate peptides that do not require post-translational modifications from the secretory pathway, such as disulfide bond formation, *N*- and *O*-glycosylation or exoprotease activity.

The role of Cym1p cytosolic activity in intracellular peptide synthesis is shown in the biosynthesis of CCK-22 in wild type cells compared to the isogenic strain with a *CYM1* disruption, which shows a significant increase in the amount of CCK-22 (Fig. 8C). Synthesis of the peptide of interest should be performed in such a way that translocation into the endoplasmatic reticulum (ER) is avoided. This can be performed either by removal of the hydrophobic amino-terminal signal sequence from proteins that enter the ER post-translationally, or by expression in a temperature sensitive secretory mutant such as *sec61*, which abolishes translocation of secretory peptides into the ER when the temperature is elevated to 37°C.

The propeptide or prepropeptide of interest will then be cytosolicly located and a potential substrate for Cym1p. Release of the peptide from its precursor will be carried out by the Cym1p activity by introduction of the cleavage site seen from proCCK, which results in the release of Gly-extended CCK-22 after endoproteolytic processing C-terminal to Lys^{61} (Ser-Ile-Val-Lys^{61} ↓) (Fig. 13A). If the peptide of interest is GLP1, synthesis can be performed as a fusion to a Cym1p cleavage site, which could be a part of proCCK (Fig. 13B). The peptide of interest will then accumulate in the cytosol and can be purified from sedimented cells after lysis.

Discussion

The secreted polypeptides varies with the growth conditions, the fraction of CCK-22 increasing when the culture reaches stationary phase, while the intracellular processed fraction remains unaltered under stress conditions. The increase in extracellular cleavage to CCK-22 as the cells enter stationary phase could indicate that extracellular endoproteases with the ability to process proCCK to CCK-22 are secreted or expressed on the cell membrane. It is known that the aspartyl proteases, Yps1p and Yps2p, exhibit cell surface activity (Komano et al., 1998). In addition, it has previously been shown that heterologous peptide expression in a *yps1* strain improved the recovery of proteins and peptides like albumin, glucagon, GLP1, GLP2 and CART by inhibiting proteolysis C-terminal to mono-basic residues (Egel-Mitani et al., 2000; Kerry-Williams et al., 1998). Thus, recent studies ((Egel-Mitani et al., 2000; Kerry-Williams et al., 1998) and those of the present inventors) show the importance of collecting secreted peptides during exponential growth in order to avoid additional extracellular processing.

ProCCK expressed in a vacuole protease deficient strain showed 30% intracellular processing at Lys^{61} in proCCK. The fraction of extracellular Lys^{61}-processing is, however, decreased to 2/3 of the observed within intact yeast cells, which reveals an intracellular degradation of CCK-22 prior to secretion. The increase in extracellular proteolysis under limited nutrient resources is probably due to an activation or upregulation in transcription of the extracellular proteases under limited nutrient resources as seen with the upregulation of *YPS1* transcription during stationary phase (Gasch et al., 2000). Part of the cell surface activity can be assigned to the yapsins, Yps1p, Yps2p and Yps3p, but some extracellular activity was sustained even in the triple mutant.

In the present study the inventors have shown that deletion of *KEX2* causes a 5 fold reduction in both the intracellular and extracellular Lys^{61}-cleavage. The *kex2* strain expressing proCCK do not only alter the cleavage of Lys^{61} in proCCK, it also changes the intracellular retention time of CCK as the intracellular concentration of CCK peptides is reduced with more than 60%, while the extracellular CCK concentration is increased almost 60% compared to wild type yeast. Moreover, analysis of the secreted CCK peptides from the *kex1 kex2* double mutant and the *kex2* mutant showed disappearance of the Tyr^{45}-Val^{60} degradation product. Thus, the removal of Lys^{61} by Kex1p was abolished in a *kex2* strain indicating an enhanced secretion rate through the *trans*-Golgi network. These results and the observations on the rapid secretion of proCCK suggest that it may be the intracellular retention caused by Kex2p that leads to

an increased synthesis of CCK-22 in wild type yeast by Yps1p and probably to some extent by Kex2p.

The type of protease responsible for the intracellular maturation of CCK-22 was investigated in an *in vitro* protease assay using a crude extract of *S. cerevisiae* to analyse the processing of synthetic human CCK-33 to CCK-22 in the presence of different inhibitors. By not including detergents in extraction of protease activity, activity from Kex2p as well as the GPI-anchored yapsins was avoided (Azaryan et al., 1993; Fuller et al., 1989; Komano et al., 1999). Of the inhibitors tested, the proteolysis was only inhibited by EDTA and 1,10 ortho-phenanthroline, and the activity could be restored by addition of the divalent cations Zn^{2+}, Co^{2+} and Mn^{2+}. This indicated that a metalloprotease participate in the maturation of CCK-22.

None of the candidate metalloproteases contain an obvious signal peptide to direct the protein into ER. Therefore, the inventors investigated strains deficient in each of the metalloprotease with the exception of mitochondrial proteases. Expression of proCCK in each of the strains resulted in unaltered maturation of CCK-22 similar to that seen in wild type yeast. However, by using the *in vitro* protease assay the inventors identified Cym1p as an endoprotease performing post-Lys cleavage of CCK-33. That Cym1p can cleave Lys^{61} in proCCK was verified by overexpression studies, showing a several fold increase in enzyme activity.

Intracellular synthesis of CCK-22 was decreased in a *cym1* strain accompanied by an increased concentration of total proCCK. In contrast, the fraction of extracellular CCK-22 was increased compared to wild type yeast with a parallel increase in total CCK. These findings are in accordance with a cytosolic location of the Cym1p activity like most insulin-degrading enzymes (Bai et al., 1996) and shows that it acts on the preproMfα1p-proCCK construct prior to translocation into the endoplasmatic reticulum. Thus, the pre-translocational degradation of proCCK is decreased by *CYM1* disruption and the total production increased.

All publications discussed above are incorporated herein in their entirety.

It will be appreciated by persons skilled in the art that numerous variations and/or modifications may be made to the invention as shown in the specific embodiments without departing from the spirit or scope of the invention as broadly described. The present embodiments are, therefore, to be considered in all respects as illustrative and not restrictive.

References

Adames, N., Blundell, K., Ashby, M.N., and Boone, C. (1995). Science *270*:464-67.

Azaryan, A.V., Wong, M., Friedman, T.C., Cawley, N.X., Estivariz, F.E., Chen, H.C., and Loh, Y.P. (1993). J. Biol. Chem. *268*:11968-75.

Bai, J.P., Hong, H.J., Rothenberger, D.A., Wong, W.D., and Buls, J.G. (1996). J. Pharm. Pharmacol. *48*:1180-84.

Brachmann, C.B., Davies, A., Cost, G.J., Caputo, E., Li, J., Hieter, P., and Boeke, J.D. (1998). Yeast *14*:115-132.

Cantor, P. and Rehfeld, J.F. (1987). Clin. Chim. Acta *168*:153-58.

Egel-Mitani, M., Andersen, A.S., Diers, I., Hach, M., Thim, L., Hastrup,S., and Vad, K. (2000). Enzyme Microb. Technol. *26*:671-77.

Egel-Mitani, M., Flygenring, H.P., and Hansen, M.T. (1990). Yeast *6*:127-137.

Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001). Nature *411*, 494-98.

Fuller, R.S., Brake, A., and Thorner, J. (1989). Proc. Natl. Acad. Sci. USA. *86*:1434-38.

Gasch, A.P., Spellman, P.T., Kao, C.M., Carmel-Harel, O., Eisen, M.B., Storz, G., Botstein, D., and Brown, P.O. (2000). Mol. Biol. Cell *11*:4241-57.

Gietz, R.D., Schiestl, R.H., Willems, A.R., and Woods, R.A. (1995). Yeast *11*:355-60.

Haseloff, J. and Gerlach, W.L. (1988). Nature, *334*: 585-91.

Hilsted, L. and Rehfeld, J.F. (1986). Anal. Biochem. *152*:119-26.

Horton, R.M., Ho, S.N., Pullen, J.K., Hunt, H.D., Cai, Z., and Pease, L.R. (1993). Methods Enzymol. *217*:270-79.

Jones, E.W. (1991). Methods Enzymol. *194*:428-53.

Julius, D., Schekman, R., and Thorner, J. (1984). Cell *36*:309-18.

Kerry-Williams, S.M., Gilbert, S.C., Evans, L.R., and Balance, D.J. (1998). Yeast *14*:161-69.

Komano, H., Rockwell, N., Wang, G.T., Krafft, G.A., and Fuller, R.S. (1999). J. Biol. Chem. *274*:24431-37.

Komano, H., Seeger, M., Gandy, S., Wang, G.T., Krafft, G.A., and Fuller, R.S. (1998). J. Biol. Chem. *273*:31648-51.

Needleman, S.B. and Wunsch, C.D. (1970). J. Mol. Biol. *48*: 443-53.

Paloheimo, L.I. and Rehfeld, J.F. (1994). Clin. Chim. Acta *229*:49-65.

Philippsen, P., Stotz, A., and Scherf, C. (1991). Methods Enzymol. *194*:169-82.

Rawlings, N.D. and Barrett, A.J. (1995). Meth. Enzymol. *248*:183-228.

Rothstein, R. (1991). Methods Enzymol. *194*:281-301.

Rourke, I.J., Johnsen, A.H., Din, N., Petersen, J.G.L., and Rehfeld, J.F. (1997). J. Biol. Chem. *272*:9720-27.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Schmidt, W.K., Tam, A., and Michaelis, S. (2000). J. Biol. Chem. *275*:6227-33.

Shippy, R., Lockner, R., Farnsworth, M. and Hampel, A. (1999). Mol. Biotech. *12*: 117-29.

Waterhouse, P.M., Graham, M.W., and Wang, M.-B. (1998). Proc Natl Acad Sci U S A, *95*: 13959-64.

Modtaget PVS

20 SEP. 2002

Claims

1. A method for producing a protein of interest in a host cell, wherein said host cell has been genetically modified in order to express significantly reduced levels of a metalloprotease which is at least 80% identical to the SEQ ID NO: 4 as compared to a parental cell, when cultured under identical conditions, comprising

a) introducing into the host cell a nucleic acid sequence encoding the protein of interest;

b) cultivating the host cell of step (a) in a suitable growth medium for production of the protein of interest and

c) isolating the protein of interst.

2. The method according to claim 1, in which the host cell is a prokaryotic cell.

3. The method according to claim 1, in which the host cell is a eukaryotic cell.

4. The method according to claim 3, in which the host cell is a non-filamentous fungal cell.

5. The method according to claim 3, in which the host cell is a filamentous fungal cell.

6. The method according to claim 4, in which the host cell is a strain of *Saccharomycces*.

7. The method according to claim 6, in which the host cell is *Saccharomyces cerevisiae*.

8. The method according to any of the proceeding claims, wherein the host cell has been genetically modified by a method selected form the group comprising gene knock-out, gene disruption, random or site directed mutagenesis, introduction of dominant-negative metalloproteases, RNA interference (RNAi) using dsRNA, catalytic nucleic acids (such as ribozymes and DNAzymes), antisense nucleic acids or a combination thereof.

9. The method according to any of the proceeding claims, wherein the level of expressed metalloprotease in the host cell is reduced more than about 50%, such as more than about 80%, such as more than about 90%, or such as more than about 95%.

10. The method according to any of the proceeding claims, wherein the host cell is essentially free of any metalloprotease activity.

11. The method according to any of the proceeding claims, in which the protein of interest is a eucaryotic protein, selected from the group comprising insulin, growth hormone, glucagon, somatostatin, inteferon, PDGF, factor VII, factor Viii, urokinase, EPO, chymosin, tissue plasminogen activator, or serum albumin.

12. The method according to any of the proceeding claims, in which the protein of interest is a protein of fungal origin, selected from the group comprising an amylolytic enzyme, an alpha-amylase, a beta-amylase, a glycoamylase, a alpha-galactosidase, a cellulytic enzyme, a lipolytic enzyme, a xylanolytic enzyme, a proteolytic enzyme, an oxidoreductase, a peroxidase, a laccase, a pectinase, or a cutinase.

13. The method according to any of the proceeding claims, in which the protein of interst is a bacterial protein, selected from the group comprising an amylolytic enzyme, an alpha-amylase, a beta-amylase, a glycoamylase, a beta-galactosidase, a cellulytic enzyme, a lipolytic enzyme, a xylanolytic enzyme, a proteolytic enzyme, an oxidoreductase, a peroxidase, a laccase, a pectinase, or a cutinase.

14. The method according to any of the proceeding claims, in which the protein of interest is a precursor, i.e. a zymogen, a hybrid protein, a protein obtained as a pro sequence or pre-pro sequence, or in unmaturated form.

15. A host cell useful for the expression of a protein of interest, wherein said cell has been genetically modified in order to express significantly reduced levels of a metalloprotease which is at least 80% identical to the SEQ ID NO: 4, as compared to a parental cell.

16. A host cell according to claim 15, wherein said cell is a prokaryotic cell.

17. A host cell according to claim 16, wherein said cell is a eukaryotic cell.

18. A host cell according to claim 17, wherein said cell is a non-filamentous fungal cell.

19. A host cell according to claim 18, wherein said cell is a filamentous fungal cell.

20. A host cell according to claim 19, wherein said host cell is a strain of *Saccharomycces*.

21. A host cell according to claim 20, wherein said host cell is *Saccharomyces cerevisiae*.

22. A host cell according to any of claims 15-21,wherein said cell has been genetically modified by a method selected from the group comprising gene knock-out, gene disruption, random or site directed mutagenesis, introduction of dominant-negative metalloproteases, RNA interference (RNAi) using dsRNA, catalytic nucleic acids (such as ribozymes and DNAzymes), antisense nucleic acids or a combination thereof.

23. A host cell according to any of claims 15-22, wherein the level of expressed metalloprotease in said cell is reduced more than about 50%, such as more than about 80%, such as more than about 90%, such as more than about 95%.

24. A host cell according to any of claims 15-23, wherein said cell is essentially free of any metalloprotease activity.

25. A host cell according to any of claims 15-24, wherein the protein of interest is an eucaryotic protein, selected from the group comprising insulin, growth hormone, glucagon, somatostatin, inteferon, PDGF, factor VII, factor Viii, urokinase, EPO, chymosin, tissue plasminogen activator, or serum albumin.

26. A host cell according to any of claims 15-25, wherein the protein of interest is a protein of fungal origin, selected from the group comprising an amylolytic enzyme, an alpha-amylase, a beta-amylase, a glycoamylase, a alpha-galactosidase, a cellulytic enzyme, a lipolytic enzyme, a xylanolytic enzyme, a proteolytic enzyme, an oxidoreductase, a peroxidase, a laccase, a pectinase, or a cutinase.

27. A host cell according to any of claims 15-26, wherein the protein of interest is a bacterial protein, selected from the group comprising an amylolytic enzyme, an alpha-amylase, a beta-amylase, a glycoamylase, a beta-galactosidase, a cellulytic enzyme, a lipolytic enzyme, a xylanolytic enzyme, a proteolytic enzyme, an oxidoreductase, a peroxidase, a laccase, a pectinase, or a cutinase.

28. A host cell according to any of claims 15-27, wherein the protein of interest is a precursor, i.e. a zymogen, a hybrid protein, a protein obtained as a pro sequence or pre-pro sequence, or in unmaturated form.

Modtaget PVS
20 SEP. 2002

SEQUENCE LISTING

<110> Rigshospitalet

<120> Methods for increasing the production of a recombinant polypeptide from a host cell

<130> 500770

<160> 51

<170> PatentIn version 3.1

<210> 1
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Pitrilysin consensus sequence

<220>
<221> MISC_FEATURE
<222> (2)..(2)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (3)..(3)
<223> X = any amino acid

<400> 1

His Xaa Xaa Glu His
1 5

<210> 2
<211> 46
<212> PRT
<213> Artificial Sequence

<220>
<223> Pitrilysin consensus sequence

<220>
<221> MISC_FEATURE
<222> (2)..(2)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (3)..(3)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (5)..(5)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (6)..(6)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (9)..(9)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (10)..(10)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (11)..(11)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (12)..(12)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (14)..(14)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (15)..(15)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (16)..(16)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (17)..(17)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (18)..(18)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (19)..(19)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (20)..(20)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (21)..(21)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (22)..(22)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (23)..(23)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (24)..(24)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (25)..(25)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (26)..(26)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (27)..(27)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (28)..(28)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (29)..(29)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (30)..(30)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (31)..(31)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (32)..(32)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (33)..(33)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (34)..(34)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (35)..(35)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (36)..(36)
<223> X = any amino acid or is absent

<220>
<221> MISC_FEATURE
<222> (37)..(37)
<223> X = any amino acid or is absent

<220>
<221> MISC_FEATURE
<222> (40)..(40)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (42)..(42)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (43)..(43)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (44)..(44)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (45)..(45)
<223> X = any amino acid

<400> 2

Gly Xaa Xaa His Xaa Xaa Glu His Xaa Xaa Xaa Xaa Gly Xaa Xaa Xaa
1 5 10 15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
20 25 30

Xaa Xaa Xaa Xaa Xaa Asn Ala Xaa Thr Xaa Xaa Xaa Xaa Thr
35 40 45

<210> 3
<211> 46
<212> PRT
<213> Artificial Sequence

<220>
<223> Pitrilysin consensus sequence

<220>
<221> MISC_FEATURE
<222> (2)..(2)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (3)..(3)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (5)..(5)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (6)..(6)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (9)..(9)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (10)..(10)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (11)..(11)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (12)..(12)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (14)..(14)
<223> X = S or T

<220>
<221> MISC_FEATURE
<222> (15)..(15)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (19)..(19)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (20)..(20)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (21)..(21)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (22)..(22)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (23)..(23)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (24)..(24)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (25)..(25)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (26)..(26)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (27)..(27)
<223> X = any amino acid or is absent

<220>
<221> MISC_FEATURE
<222> (28)..(28)
<223> X = any amino acid or is absent

<220>
<221> MISC_FEATURE
<222> (30)..(30)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (31)..(31)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (32)..(32)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (33)..(33)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (34)..(34)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (35)..(35)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (36)..(36)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (37)..(37)
<223> X = any amino acid or is absent

<220>
<221> MISC_FEATURE
<222> (40)..(40)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (42)..(42)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (43)..(43)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (44)..(44)
<223> X = D or E

<220>
<221> MISC_FEATURE
<222> (45)..(45)
<223> X = any amino acid

<400> 3

Gly Xaa Xaa His Xaa Xaa Glu His Xaa Xaa Xaa Xaa Gly Xaa Xaa Lys
1 5 10 15

Tyr Pro Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Xaa Xaa Xaa
20 25 30

Xaa Xaa Xaa Xaa Xaa Asn Ala Xaa Thr Xaa Xaa Xaa Xaa Thr
35 40 45

<210> 4
<211> 989
<212> PRT
<213> Saccharomyces cerevisiae

<400> 4

Met Leu Arg Phe Gln Arg Phe Ala Ser Ser Tyr Ala Gln Ala Gln Ala
1 5 10 15

Val Arg Lys Tyr Pro Val Gly Gly Ile Phe His Gly Tyr Glu Val Arg
20 25 30

Arg Ile Leu Pro Val Pro Glu Leu Arg Leu Thr Ala Val Asp Leu Val
35 40 45

His Ser Gln Thr Gly Ala Glu His Leu His Ile Asp Arg Asp Asp Lys
50 55 60

Asn Asn Val Phe Ser Ile Ala Phe Lys Thr Asn Pro Pro Asp Ser Thr
65 70 75 80

Gly Val Pro His Ile Leu Glu His Thr Thr Leu Cys Gly Ser Val Lys
85 90 95

```
Tyr Pro Val Arg Asp Pro Phe Phe Lys Met Leu Asn Lys Ser Leu Ala
            100                 105                 110

Asn Phe Met Asn Ala Met Thr Gly Pro Asp Tyr Thr Phe Phe Pro Phe
        115                 120                 125

Ser Thr Thr Asn Pro Gln Asp Phe Ala Asn Leu Arg Gly Val Tyr Leu
    130                 135                 140

Asp Ser Thr Leu Asn Pro Leu Leu Lys Gln Glu Asp Phe Asp Gln Glu
145                 150                 155                 160

Gly Trp Arg Leu Glu His Lys Asn Ile Thr Asp Pro Glu Ser Asn Ile
                165                 170                 175

Val Phe Lys Gly Val Val Tyr Asn Glu Met Lys Gly Gln Ile Ser Asn
            180                 185                 190

Ala Asn Tyr Tyr Phe Trp Ser Lys Phe Gln Gln Ser Ile Tyr Pro Ser
        195                 200                 205

Leu Asn Asn Ser Gly Gly Asp Pro Met Lys Ile Thr Asp Leu Arg Tyr
    210                 215                 220

Gly Asp Leu Leu Asp Phe His His Lys Asn Tyr His Pro Ser Asn Ala
225                 230                 235                 240

Lys Thr Phe Thr Tyr Gly Asn Leu Pro Leu Val Asp Thr Leu Lys Gln
                245                 250                 255

Leu Asn Glu Gln Phe Ser Gly Tyr Gly Lys Arg Ala Arg Lys Asp Lys
            260                 265                 270

Leu Leu Met Pro Ile Asp Leu Lys Lys Asp Ile Asp Val Lys Leu Leu
        275                 280                 285

Gly Gln Ile Asp Thr Met Leu Pro Pro Glu Lys Gln Thr Lys Ala Ser
    290                 295                 300

Met Thr Trp Ile Cys Gly Ala Pro Gln Asp Thr Tyr Asp Thr Phe Leu
305                 310                 315                 320

Leu Lys Val Leu Gly Asn Leu Leu Met Asp Gly His Ser Ser Val Met
                325                 330                 335
```

```
Tyr Gln Lys Leu Ile Glu Ser Gly Ile Gly Leu Glu Phe Ser Val Asn
            340                 345                 350

Ser Gly Val Glu Pro Thr Thr Ala Val Asn Leu Leu Thr Val Gly Ile
        355                 360                 365

Gln Gly Val Ser Asp Ile Glu Ile Phe Lys Asp Thr Val Asn Asn Ile
    370                 375                 380

Phe Gln Asn Leu Leu Glu Thr Glu His Pro Phe Asp Arg Lys Arg Ile
385                 390                 395                 400

Asp Ala Ile Ile Glu Gln Leu Glu Leu Ser Lys Lys Asp Gln Lys Ala
                405                 410                 415

Asp Phe Gly Leu Gln Leu Leu Tyr Ser Ile Leu Pro Gly Trp Thr Asn
            420                 425                 430

Lys Ile Asp Pro Phe Glu Ser Leu Leu Phe Glu Asp Val Leu Gln Arg
        435                 440                 445

Phe Arg Gly Asp Leu Glu Thr Lys Gly Asp Thr Leu Phe Gln Asp Leu
    450                 455                 460

Ile Arg Lys Tyr Ile Val His Lys Pro Cys Phe Thr Phe Ser Ile Gln
465                 470                 475                 480

Gly Ser Glu Glu Phe Ser Lys Ser Leu Asp Asp Glu Glu Gln Thr Arg
                485                 490                 495

Leu Arg Glu Lys Ile Thr Ala Leu Asp Glu Gln Asp Lys Lys Asn Ile
            500                 505                 510

Phe Lys Arg Gly Ile Leu Leu Gln Glu Lys Gln Asn Glu Lys Glu Asp
        515                 520                 525

Leu Ser Cys Leu Pro Thr Leu Gln Ile Lys Asp Ile Pro Arg Ala Gly
    530                 535                 540

Asp Lys Tyr Ser Ile Glu Gln Lys Asn Asn Thr Met Ser Arg Ile Thr
545                 550                 555                 560

Asp Thr Asn Gly Ile Thr Tyr Val Arg Gly Lys Arg Leu Leu Asn Asp
                565                 570                 575
```

```
Ile Ile Pro Phe Glu Leu Phe Pro Tyr Leu Pro Leu Phe Ala Glu Ser
            580                 585                 590

Leu Thr Asn Leu Gly Thr Thr Thr Glu Ser Phe Ser Glu Ile Glu Asp
        595                 600                 605

Gln Ile Lys Leu His Thr Gly Gly Ile Ser Thr His Val Glu Val Thr
    610                 615                 620

Ser Asp Pro Asn Thr Thr Glu Pro Arg Leu Ile Phe Gly Phe Asp Gly
625                 630                 635                 640

Trp Ser Leu Asn Ser Lys Thr Asp His Ile Phe Glu Phe Trp Ser Lys
                645                 650                 655

Ile Leu Leu Glu Thr Asp Phe His Lys Asn Ser Asp Lys Leu Lys Val
            660                 665                 670

Leu Ile Arg Leu Leu Ala Ser Ser Asn Thr Ser Ser Val Ala Asp Ala
        675                 680                 685

Gly His Ala Phe Ala Arg Gly Tyr Ser Ala Ala His Tyr Arg Ser Ser
    690                 695                 700

Gly Ala Ile Asn Glu Thr Leu Asn Gly Ile Glu Gln Leu Gln Phe Ile
705                 710                 715                 720

Asn Arg Leu His Ser Leu Leu Asp Asn Glu Glu Thr Phe Gln Arg Glu
                725                 730                 735

Val Val Asp Lys Leu Thr Glu Leu Gln Lys Tyr Ile Val Asp Thr Asn
            740                 745                 750

Asn Met Asn Phe Phe Ile Thr Ser Asp Ser Asp Val Gln Ala Lys Thr
        755                 760                 765

Val Glu Ser Gln Ile Ser Lys Phe Met Glu Arg Leu Pro His Gly Ser
    770                 775                 780

Cys Leu Pro Asn Gly Pro Lys Thr Ser Asp Tyr Pro Leu Ile Gly Ser
785                 790                 795                 800

Lys Cys Lys His Thr Leu Ile Lys Phe Pro Phe Gln Val His Tyr Thr
                805                 810                 815
```

```
Ser Gln Ala Leu Leu Gly Val Pro Tyr Thr His Lys Asp Gly Ser Ala
            820                 825                 830

Leu Gln Val Met Ser Asn Met Leu Thr Phe Lys His Leu His Arg Glu
        835                 840                 845

Val Arg Glu Lys Gly Gly Ala Tyr Gly Gly Gly Ala Ser Tyr Ser Ala
    850                 855                 860

Leu Ala Gly Ile Phe Ser Phe Tyr Ser Tyr Arg Asp Pro Gln Pro Leu
865                 870                 875                 880

Lys Ser Leu Glu Thr Phe Lys Asn Ser Gly Arg Tyr Ile Leu Asn Asp
                885                 890                 895

Ala Lys Trp Gly Val Thr Asp Leu Asp Glu Ala Lys Leu Thr Ile Phe
            900                 905                 910

Gln Gln Val Asp Ala Pro Lys Ser Pro Lys Gly Glu Gly Val Thr Tyr
        915                 920                 925

Phe Met Ser Gly Val Thr Asp Asp Met Lys Gln Ala Arg Arg Glu Gln
    930                 935                 940

Leu Leu Asp Val Ser Leu Leu Asp Val His Arg Val Ala Glu Lys Tyr
945                 950                 955                 960

Leu Leu Asn Lys Glu Gly Val Ser Thr Val Ile Gly Pro Gly Ile Glu
                965                 970                 975

Gly Lys Thr Val Ser Pro Asn Trp Glu Val Lys Glu Leu
            980                 985
```

<210> 5
<211> 882
<212> PRT
<213> Schizosaccharomyces pombe

<400> 5

```
Met Asn Tyr Ala Lys Leu Ser Ile Ala Phe Ser Lys Lys Thr Ile Lys
1               5                   10                  15

Thr His Asn Cys Arg Leu Phe Gln Arg Trp Leu His Val Gly Asp Lys
            20                  25                  30
```

```
Val His Asp Phe Arg Val Val Asp Thr Lys Lys Val Pro Glu Leu Gln
        35                  40                  45

Leu Asn Tyr Thr Arg Leu Lys His Glu Pro Thr Asn Ala Asp Met Ile
    50                  55                  60

His Leu Asp Arg Glu Asp Pro Asn Ser Val Phe Ser Ile Gly Phe Gln
65                  70                  75                  80

Thr Pro Ala Glu Asn Asp Glu Gly Ile Pro His Ile Leu Glu His Thr
                85                  90                  95

Thr Leu Cys Gly Ser Asn Lys Tyr Pro Val Arg Asp Pro Phe Phe Lys
            100                 105                 110

Met Leu Asn Arg Ser Leu Ala Thr Phe Met Asn Ala Phe Thr Ala Ser
        115                 120                 125

Asp Phe Thr Phe Tyr Pro Phe Ala Thr Val Asn Thr Thr Asp Tyr Lys
    130                 135                 140

Asn Leu Arg Asp Val Tyr Leu Asp Ala Thr Leu Phe Pro Lys Leu Arg
145                 150                 155                 160

Lys Leu Asp Phe Leu Gln Glu Gly Trp Arg Phe Glu His Ala Asp Val
                165                 170                 175

Asn Asp Lys Lys Ser Pro Ile Ile Phe Asn Gly Val Val Tyr Asn Glu
            180                 185                 190

Met Lys Gly Gln Val Ser Asp Ser Ser Tyr Ile Phe Tyr Met Leu Phe
        195                 200                 205

Gln Gln His Leu Phe Gln Gly Thr Ala Tyr Gly Phe Asn Ser Gly Gly
    210                 215                 220

Asp Pro Leu Ala Ile Pro Asp Leu Lys Tyr Glu Glu Leu Val Lys Phe
225                 230                 235                 240

His Arg Ser His Tyr His Pro Ser Asn Ala Lys Ile Leu Ser Tyr Gly
                245                 250                 255

Ser Phe Pro Leu Glu Asp Asn Leu Ser Ala Leu Ser Glu Thr Phe Arg
            260                 265                 270
```

```
Pro Phe Ser Lys Arg Glu Leu Asn Leu Pro Asn Thr Phe Leu Lys Glu
        275                 280                 285

Phe Asp Gln Glu Lys Arg Val Val Glu Tyr Gly Pro Leu Asp Pro Val
    290                 295                 300

Met Ala Pro Gly Arg Gln Val Lys Thr Ser Ile Ser Phe Leu Ala Asn
305                 310                 315                 320

Asp Thr Ser Asn Val Tyr Glu Thr Phe Ala Leu Lys Val Leu Ser Lys
                325                 330                 335

Leu Cys Phe Asp Gly Phe Ser Ser Pro Phe Tyr Lys Ala Leu Ile Glu
            340                 345                 350

Ser Gly Leu Gly Thr Asp Phe Ala Pro Asn Ser Gly Tyr Asp Ser Thr
        355                 360                 365

Thr Lys Arg Gly Ile Phe Ser Val Gly Leu Glu Gly Ala Ser Glu Glu
    370                 375                 380

Ser Leu Ala Lys Ile Glu Asn Leu Val Tyr Ser Ile Phe Asn Asp Leu
385                 390                 395                 400

Ala Leu Lys Gly Phe Glu Asn Glu Lys Leu Glu Ala Ile Leu His Gln
                405                 410                 415

Met Glu Ile Ser Leu Lys His Lys Ser Ala His Phe Gly Ile Gly Leu
            420                 425                 430

Ala Gln Ser Leu Pro Phe Asn Trp Phe Asn Gly Ala Asp Pro Ala Asp
        435                 440                 445

Trp Leu Ser Phe Asn Lys Gln Ile Glu Trp Leu Lys Gln Lys Asn Ser
    450                 455                 460

Asp Gly Lys Leu Phe Gln Lys Leu Ile Lys Lys Tyr Ile Leu Glu Asn
465                 470                 475                 480

Lys Ser Arg Phe Val Phe Thr Met Leu Pro Ser Ser Thr Phe Pro Gln
                485                 490                 495

Arg Leu Gln Glu Ala Glu Ala Lys Lys Leu Gln Glu Arg Thr Ser Lys
            500                 505                 510
```

```
Leu Thr Asp Glu Asp Ile Ala Glu Ile Glu Lys Thr Ser Val Lys Leu
        515                 520                 525

Leu Glu Ala Gln Ser Thr Pro Ala Asp Thr Ser Cys Leu Pro Thr Leu
    530                 535                 540

Ser Val Ser Asp Ile Pro Glu Thr Ile Asp Glu Thr Lys Leu Lys Phe
545                 550                 555                 560

Leu Asp Ile Ala Gly Met Lys Ala Gln Trp Tyr Asp Leu Ala Ala Gly
                565                 570                 575

Leu Thr Tyr Ile Arg Leu Leu Leu Pro Leu Lys Asn Phe Pro Glu Ser
            580                 585                 590

Leu Ile Pro Tyr Leu Pro Val Tyr Cys Asp Ala Cys Leu Asn Leu Gly
        595                 600                 605

Thr His Ser Glu Ser Ile Gly Asp Leu Glu His Gln Ile Arg Arg Tyr
    610                 615                 620

Thr Gly Gly Ile Ser Ile Ser Pro Ser Ala Val Thr Asn Asn Ser Asp
625                 630                 635                 640

Val Ser Lys Tyr Glu Leu Gly Ile Ala Ile Ser Gly Tyr Ala Leu Asp
                645                 650                 655

Lys Asn Val Gly Lys Leu Val Glu Leu Ile Asn Lys Ala Phe Trp Asn
            660                 665                 670

Thr Asn Leu Ser Asn Thr Asp Lys Leu Ala Ile Met Leu Lys Thr Ser
        675                 680                 685

Val Ser Gly Ile Thr Asp Gly Ile Ala Glu Lys Gly His Ser Phe Ala
    690                 695                 700

Lys Val Ser Ser Ala Ser Gly Leu Thr Glu Lys Thr Ser Ile Thr Glu
705                 710                 715                 720

Gln Leu Gly Gly Leu Thr Gln Val Lys Leu Leu Ser Gln Leu Ser Arg
                725                 730                 735

Glu Glu Ser Phe Gly Pro Leu Val Glu Lys Leu Thr Ala Ile Arg Glu
            740                 745                 750
```

```
Ile Leu Arg Gly Thr Ser Gly Phe Lys Ala Ala Ile Asn Ala Ser Pro
        755                 760                 765

Thr Gln His Glu Val Val Glu Lys Ala Leu Gln Lys Phe Met Lys Ser
    770                 775                 780

Arg Gly Val Asn Gln Gln Thr Gln Thr Lys Ser Thr Ser Lys Glu Arg
785                 790                 795                 800

Asn Gly Ile Asn Ser Ile Lys Thr Tyr His Glu Leu Pro Phe Gln Thr
                805                 810                 815

Tyr Phe Ala Ala Lys Ser Cys Leu Gly Val Pro Tyr Thr His Pro Asp
            820                 825                 830

Gly Ala Pro Leu Gln Ile Leu Ser Ser Leu Leu Thr His Lys Tyr Leu
        835                 840                 845

His Gly Glu Ile Arg Glu Lys Gly Gly Ala Tyr Gly Ala Gly Leu Ser
    850                 855                 860

Tyr Ser Gly Ile Asp Gly Val Leu Ser Phe Phe Thr Tyr Arg Asp Ser
865                 870                 875                 880

Asp Pro
```

<210> 6
<211> 973
<212> PRT
<213> Clostridium perfringens

<400> 6

```
Met Asn Phe Lys Glu Asn Asn Ile Tyr Ser Gly Phe Lys Leu Leu Asn
1               5                   10                  15

Ile Glu Asn Leu Asn Glu Ile Gly Gly Val Gly Leu Arg Phe Glu His
            20                  25                  30

Glu Lys Thr Lys Ala Lys Leu Ile Lys Ile Leu Ser Glu Asp Asp Asn
        35                  40                  45

Lys Cys Phe Ala Ile Gly Phe Arg Thr Pro Pro Glu Asn Ser Thr Gly
    50                  55                  60
```

```
Val Pro His Ile Leu Glu His Ser Val Leu Cys Gly Ser Arg Lys Phe
65                  70                  75                  80

Asn Thr Lys Glu Pro Phe Val Glu Leu Leu Lys Gly Ser Leu Asn Thr
                85                  90                  95

Phe Leu Asn Ala Met Thr Tyr Pro Asp Lys Thr Ile Tyr Pro Val Ala
            100                 105                 110

Ser Arg Asn Glu Lys Asp Phe Met Asn Leu Met Asp Val Tyr Leu Asp
        115                 120                 125

Ala Val Leu Tyr Pro Asn Ile Tyr Lys His Lys Glu Ile Phe Met Gln
    130                 135                 140

Glu Gly Trp His Tyr Tyr Ile Glu Asn Lys Glu Asp Glu Leu Lys Tyr
145                 150                 155                 160

Asn Gly Val Val Tyr Asn Glu Met Lys Gly Ala Tyr Ser Ser Pro Asp
                165                 170                 175

Ser Ile Leu Tyr Arg Lys Ile Pro Gln Thr Ile Tyr Pro Asp Thr Cys
            180                 185                 190

Tyr Ala Leu Ser Ser Gly Gly Asp Pro Asp Glu Ile Pro Asn Leu Thr
        195                 200                 205

Tyr Glu Glu Phe Val Glu Phe His Lys Lys Tyr Tyr His Pro Ser Asn
    210                 215                 220

Ser Tyr Ile Phe Leu Tyr Gly Asn Gly Asp Thr Glu Lys Glu Leu Glu
225                 230                 235                 240

Phe Ile Asn Glu Glu Tyr Leu Lys Asn Phe Glu Tyr Lys Glu Ile Asp
                245                 250                 255

Ser Glu Ile Lys Glu Gln Lys Ser Phe Glu Ser Met Lys Glu Glu Ser
            260                 265                 270

Phe Thr Tyr Gly Ile Ala Glu Ser Glu Asp Leu Asn His Lys Ser Tyr
        275                 280                 285

Tyr Ser Leu Asn Phe Val Ile Gly Asp Ala Thr Asp Gly Glu Lys Gly
    290                 295                 300
```

```
Leu Ala Phe Asp Val Leu Ala Tyr Leu Leu Thr Arg Ser Thr Ala Ala
305                 310                 315                 320

Pro Leu Lys Lys Ala Leu Ile Asp Ala Gly Ile Gly Lys Ala Val Ser
                325                 330                 335

Gly Asp Phe Asp Asn Ser Thr Lys Gln Ser Ala Phe Thr Val Leu Val
            340                 345                 350

Lys Asn Ala Glu Leu Asn Lys Glu Glu Glu Phe Lys Lys Val Val Met
        355                 360                 365

Asp Thr Leu Lys Asp Leu Val Glu Asn Gly Ile Asp Lys Glu Leu Ile
    370                 375                 380

Glu Ala Ser Ile Asn Arg Val Glu Phe Glu Leu Arg Glu Gly Asp Tyr
385                 390                 395                 400

Gly Ser Tyr Pro Asn Gly Leu Ile Tyr Tyr Leu Lys Val Met Asp Ser
                405                 410                 415

Trp Leu Tyr Asp Gly Asp Pro Tyr Val His Leu Glu Tyr Glu Lys Asn
            420                 425                 430

Leu Glu Lys Ile Lys Ser Ala Leu Thr Ser Asn Tyr Phe Glu Asp Leu
        435                 440                 445

Ile Glu Arg Tyr Met Ile Asn Asn Thr His Ser Ser Leu Val Ser Leu
    450                 455                 460

His Pro Glu Lys Gly Ile Asn Glu Lys Lys Ser Ala Glu Leu Lys Lys
465                 470                 475                 480

Lys Leu Glu Glu Ile Lys Asn Ser Phe Asp Glu Lys Thr Leu Asn Glu
                485                 490                 495

Ile Ile Asp Asn Cys Lys Lys Leu Lys Glu Arg Gln Ser Thr Pro Asp
            500                 505                 510

Lys Lys Glu Asp Leu Glu Ser Ile Pro Met Leu Ser Leu Glu Asp Ile
        515                 520                 525

Asp Lys Glu Ala Thr Lys Ile Pro Thr Glu Glu Lys Glu Ile Asp Gly
    530                 535                 540
```

```
Ile Thr Thr Leu His His Asp Phe His Thr Asn Lys Ile Asp Tyr Val
545                 550                 555                 560

Asn Phe Phe Phe Asn Thr Asn Ser Val Pro Glu Asp Leu Ile Pro Tyr
                565                 570                 575

Val Gly Leu Leu Cys Asp Ile Leu Gly Lys Cys Gly Thr Glu Asn Tyr
            580                 585                 590

Asp Tyr Ser Lys Leu Ser Asn Ala Ile Asn Ile Ser Thr Gly Gly Ile
        595                 600                 605

Ser Phe Gly Ala Ile Thr Phe Ala Asn Leu Lys Lys Asn Asn Glu Phe
    610                 615                 620

Arg Pro Tyr Leu Glu Ile Ser Tyr Lys Ala Leu Ser Ser Lys Thr Asn
625                 630                 635                 640

Lys Ala Ile Glu Leu Val Asp Glu Ile Val Asn His Thr Asp Leu Asp
                645                 650                 655

Asp Met Asp Arg Ile Met Gln Ile Ile Arg Glu Lys Arg Ala Arg Leu
            660                 665                 670

Glu Gly Ala Ile Phe Asp Ser Gly His Arg Ile Ala Met Lys Lys Val
        675                 680                 685

Leu Ser Tyr Ser Thr Asn Arg Gly Ala Tyr Asp Glu Lys Ile Ser Gly
    690                 695                 700

Leu Asp Tyr Tyr Asp Phe Leu Val Asn Ile Glu Lys Glu Asp Lys Lys
705                 710                 715                 720

Ser Thr Ile Ser Asp Ser Leu Lys Lys Val Arg Asp Leu Ile Phe Asn
                725                 730                 735

Lys Gly Asn Met Leu Ile Ser Tyr Ser Gly Lys Glu Glu Glu Tyr Glu
            740                 745                 750

Asn Phe Lys Glu Lys Val Lys Tyr Leu Ile Ser Lys Thr Asn Asn Asn
        755                 760                 765

Asp Phe Glu Lys Glu Glu Tyr Asn Phe Glu Leu Gly Lys Lys Asn Glu
    770                 775                 780
```

```
Gly Leu Leu Thr Gln Gly Asn Val Gln Tyr Val Ala Lys Gly Gly Asn
785                 790                 795                 800

Tyr Lys Thr His Gly Tyr Lys Tyr Ser Gly Ala Leu Ser Leu Leu Glu
                805                 810                 815

Ser Ile Leu Gly Phe Asp Tyr Leu Trp Asn Ala Val Arg Val Lys Gly
            820                 825                 830

Gly Ala Tyr Gly Val Phe Ser Asn Phe Arg Arg Asp Gly Gly Ala Tyr
        835                 840                 845

Ile Val Ser Tyr Arg Asp Pro Asn Ile Lys Ser Thr Leu Glu Ala Tyr
    850                 855                 860

Asp Asn Ile Pro Lys Tyr Leu Asn Asp Phe Glu Ala Asp Glu Arg Glu
865                 870                 875                 880

Met Thr Lys Tyr Ile Ile Gly Thr Ile Arg Lys Tyr Asp Gln Pro Ile
                885                 890                 895

Ser Asn Gly Ile Lys Gly Asp Ile Ala Val Ser Tyr Tyr Leu Ser Asn
            900                 905                 910

Phe Thr Tyr Glu Asp Leu Gln Lys Glu Arg Glu Glu Ile Ile Asn Ala
        915                 920                 925

Asp Val Glu Lys Ile Lys Ser Phe Ala Pro Met Ile Lys Asp Leu Met
    930                 935                 940

Lys Glu Asp Tyr Ile Cys Val Leu Gly Asn Glu Glu Lys Ile Lys Glu
945                 950                 955                 960

Asn Lys Asp Leu Phe Asn Asn Ile Lys Ser Val Ile Lys
                965                 970
```

<210> 7
<211> 971
<212> PRT
<213> Borrelia burgdorferi

<400> 7

```
Met Lys Lys Lys Lys Ile Phe Lys Leu Ile Ser Lys Thr Tyr Leu Glu
1               5                   10                  15
```

```
Glu His Asp Ala Glu Gly Tyr Tyr Phe Lys His Glu Ser Gly Leu Glu
            20                  25                  30

Val Phe His Leu Lys Ser Asp Ser Phe Lys Glu Asn Ala Phe Cys Ile
        35                  40                  45

Ala Phe Lys Thr Ile Pro Ser Asn Asn Thr Gly Val Ala His Val Leu
    50                  55                  60

Glu His Thr Ile Phe Cys Gly Ser Ser Lys Tyr Lys Ile Lys Asp Pro
65                  70                  75                  80

Phe Leu Tyr Leu Leu Lys Gly Ser Leu Asn Thr Phe Leu Asn Ala Met
                85                  90                  95

Thr Phe Pro Asp Lys Thr Ile Tyr Pro Ala Ala Ser Thr Ile Glu Lys
            100                 105                 110

Asp Tyr Phe Asn Leu Phe Asn Ile Tyr Ala Asp Ser Ile Phe Asn Pro
        115                 120                 125

Leu Leu Lys Lys Glu Ser Phe Met Gln Glu Gly Tyr Asn Ile Asn Pro
    130                 135                 140

Lys Asp Phe Lys Val Ser Gly Ile Val Phe Asn Glu Met Lys Gly Ser
145                 150                 155                 160

Tyr Ser Asn Lys Asn Ser Leu Ile Asn Glu Ile Val Ser Ser Ser Leu
                165                 170                 175

Phe Glu Glu Gly Ala Tyr Lys Tyr Asp Ser Gly Gly Ile Pro Thr Asn
            180                 185                 190

Ile Ile Asp Leu Thr Tyr Glu Ser Phe Leu Asp Phe Tyr Lys Lys Tyr
        195                 200                 205

Tyr Thr Leu Glu Asn Cys Lys Ile Phe Leu Cys Gly Asn Thr Gln Thr
    210                 215                 220

Glu Lys Asn Leu Asn Phe Ile Glu Lys Tyr Ile Ile Arg Pro Tyr Lys
225                 230                 235                 240

Lys Glu Lys Ser Asn Val Asn Ile Asn Ile Glu Asn Val Lys Arg Trp
                245                 250                 255
```

```
Glu Lys Gly Lys Lys Leu Thr Tyr Lys Ile Pro Lys Glu Asn Asp Asn
            260                 265                 270

Ser Leu Gly Val Tyr Thr Ile Asn Trp Leu Cys Thr Glu Ile Asn Asn
        275                 280                 285

Ile Glu Asp Ser Ile Gly Leu Glu Ile Leu Ser Glu Ile Leu Leu Asp
    290                 295                 300

Asp Ser Cys Ser Phe Thr Ile Asn Ile Leu Lys Ser Gly Ile Gly Glu
305                 310                 315                 320

Asp Ile Ala His Ile Ser Gly Ile Asn Thr Asp Leu Lys Glu Ser Ile
                325                 330                 335

Phe Ser Phe Gly Leu Gln Asn Val Val Glu Asn Lys Glu Lys Glu Phe
            340                 345                 350

Lys Asn Leu Val Phe Ser Glu Leu Lys Asn Leu Val Lys Asn Lys Ile
        355                 360                 365

Pro Lys Glu Leu Ile Lys Gly Ile Leu Phe Gly Tyr Glu Phe Ala Leu
    370                 375                 380

Lys Glu Glu Lys Gly Gln Asn Phe Pro Ile Ala Leu Met Ile Lys Ser
385                 390                 395                 400

Phe Lys Gly Trp Leu Asn Gly Leu His Pro Ile Lys Thr Leu Gln Thr
                405                 410                 415

Ser Tyr Tyr Ile Asn Glu Ile Thr Asn Lys Leu Glu Lys Gly Ile Tyr
            420                 425                 430

Tyr Phe Glu Asn Leu Ile Glu Lys Tyr Leu Ile Phe Asn Asn His Tyr
        435                 440                 445

Thr Leu Ile Ser Phe Ile Pro Ser His Asp Thr Glu Lys Glu Met Glu
    450                 455                 460

Glu Glu Ile Glu Lys Lys Leu Met Ala Arg Glu Ile Glu Ile Lys Gln
465                 470                 475                 480

Asn Pro Glu Glu Phe Leu Gln Phe Lys Lys Asp Tyr Asn Gln Phe Lys
                485                 490                 495
```

```
Lys Tyr Gln Asn Lys Lys Asp Ser Lys Ala Asp Ile Ala Lys Leu Pro
            500                 505                 510

Leu Leu Lys Ile Glu Asp Leu Pro Lys Gln Ile Glu Lys Ser Leu Asp
        515                 520                 525

Leu Asn Glu Ile Lys Glu Leu Asn Leu His Ser Phe Lys Phe Lys Ser
    530                 535                 540

Asn Asn Ile Phe Asn Val Asn Leu Phe Phe Lys Leu Asp Phe Leu Glu
545                 550                 555                 560

Lys Glu Asp Tyr Ile Tyr Leu Ser Leu Phe Lys Arg Ala Leu Gln Asp
                565                 570                 575

Leu Ser Thr Lys Asn Tyr Ser Tyr Ile Asn Ile Asn Asn Lys Ile Gln
            580                 585                 590

Asn Thr Leu Gly Gln Ile Asn Ile Ser Glu Ser Tyr Asp Glu Asp Ile
        595                 600                 605

Asp Gly Asn Ile Leu Asn Ser Phe Asn Ile Ser Phe Lys Ser Phe Asn
    610                 615                 620

Asn Lys Val Lys Glu Ser Phe Glu Leu Ile Lys Glu Ile Leu Ile Asn
625                 630                 635                 640

Ile Asn Phe His Asp Tyr Glu Arg Leu Lys Glu Ile Thr Leu Ser Leu
                645                 650                 655

Lys Asn Asp Phe Lys Ser Leu Leu Ile Pro Lys Gly His Leu Leu Ala
            660                 665                 670

Met Leu Arg Ser Lys Ser Lys Leu Lys Leu Asn Glu Tyr Leu Lys Glu
        675                 680                 685

Leu Gln Asn Gly Ile Thr Gly Arg Glu Phe Trp Gln Lys Ala Lys Thr
    690                 695                 700

Asp Thr Glu Ser Leu Lys Glu Ile Ala Asn Lys Leu Asp Asn Leu Lys
705                 710                 715                 720

Asn Lys Ile Ile Leu Lys Asn Asn Leu Ser Ala Leu Ile Met Gly Asn
                725                 730                 735
```

```
Thr Asp Asp Ile Leu Lys Asn Leu Glu Asn Glu Phe Phe Asn Leu Lys
            740                 745                 750

Glu Ser Leu Glu Glu Ser Asn His Tyr Asn Gly Leu Leu Asn Leu Asp
        755                 760                 765

Ala Asn Ser Lys Ala Leu Arg Glu Ile Ile Ile Ile Gln Ser Lys Val
    770                 775                 780

Ala Phe Asn Ala Ile Cys Phe Pro Ser Tyr Lys Ile Asn Asp Glu Asn
785                 790                 795                 800

Tyr Pro Lys Ala Asn Phe Leu Glu His Val Leu Arg Ser Gly Ile Phe
                805                 810                 815

Trp Glu Lys Ile Arg Val Met Gly Gly Ala Tyr Gly Ala Ser Ala Ser
            820                 825                 830

Ile Ala Asn Gly Ile Phe Ser Phe Ala Ser Tyr Arg Asp Pro Asn Phe
        835                 840                 845

Thr Lys Thr Tyr Gln Ala Phe Glu Lys Ser Leu Glu Glu Leu Ala Asn
    850                 855                 860

Asn Lys Met Thr Asp Asp Glu Ile Tyr Thr Tyr Leu Ile Gly Leu Ile
865                 870                 875                 880

Gly Thr Asn Ile Tyr Val Lys Thr Lys Ala Thr Glu Ala Leu Gln Ser
                885                 890                 895

Tyr Arg Arg Lys Met Leu Asn Ile Ser Asp Ser Leu Arg Gln Asp Ile
            900                 905                 910

Arg Asn Ala Tyr Phe Thr Ile Thr Pro Gln Asp Ile Lys Glu Ile Ser
        915                 920                 925

Thr Lys Ile Leu Thr Gln Ile Arg Gln His Asn Ser Ile Ala Ser Leu
    930                 935                 940

Val Asn Asn Gln Ile Tyr Glu Glu Glu Lys Asn Asn Leu Glu Lys Leu
945                 950                 955                 960

Ile Gly Lys Glu Tyr Ser Leu Lys Lys Ile Tyr
                965                 970
```

<210> 8
<211> 995
<212> PRT
<213> Caenorhabditis elegans

<400> 8

```
Met Ser Ala Ser Lys Leu Trp Ser Cys Thr Glu Thr Val Leu Asn Gly
1               5                  10                  15

Gly Ile Lys Leu Phe Leu Tyr Ser Ser Lys Asn Thr Lys Leu Arg Val
            20                  25                  30

Ala Ile Gly Glu Val Pro Gly Pro Met Val His Gly Ala Val Ser Phe
        35                  40                  45

Val Thr Glu Ala Asp Ser Asp Asp Gly Leu Pro His Thr Leu Glu His
    50                  55                  60

Leu Val Phe Met Gly Ser Lys Lys Tyr Pro Phe Lys Gly Val Leu Asp
65                  70                  75                  80

Val Ile Ala Asn Arg Cys Leu Ala Asp Gly Thr Asn Ala Trp Thr Asp
                85                  90                  95

Thr Asp His Thr Ala Tyr Thr Leu Ser Thr Val Gly Ser Asp Gly Phe
            100                 105                 110

Leu Lys Val Leu Pro Val Tyr Ile Asn His Leu Leu Thr Pro Met Leu
        115                 120                 125

Thr Ala Ser Gln Phe Ala Thr Glu Val His His Ile Thr Gly Glu Gly
    130                 135                 140

Asn Asp Ala Gly Val Val Tyr Ser Glu Met Gln Asp His Glu Ser Glu
145                 150                 155                 160

Met Glu Ser Ile Met Asp Arg Lys Thr Lys Glu Val Ile Tyr Pro Pro
                165                 170                 175

Phe Asn Pro Tyr Ala Val Asp Thr Gly Gly Arg Leu Lys Asn Leu Arg
            180                 185                 190

Glu Ser Cys Thr Leu Glu Lys Val Arg Asp Tyr His Lys Lys Phe Tyr
        195                 200                 205
```

```
His Leu Ser Asn Met Val Val Thr Val Cys Gly Met Val Asp His Asp
    210                 215                 220

Gln Val Leu Glu Ile Met Asn Asn Val Glu Asn Glu His Met Ser Thr
225                 230                 235                 240

Val Pro Asp His Phe Pro Lys Pro Phe Ser Phe Ala Leu Ser Asp Ile
                245                 250                 255

Lys Glu Ser Thr Val His Arg Val Glu Cys Pro Thr Asp Asp Ala Ser
            260                 265                 270

Arg Gly Ala Val Glu Val Ala Trp Phe Ala His Ser Pro Ser Asp Leu
        275                 280                 285

Glu Thr His Ser Ser Leu His Val Leu Phe Asp Tyr Leu Ser Asn Thr
    290                 295                 300

Ser Val Ala Pro Leu Gln Lys Asp Phe Ile Leu Leu Glu Asp Pro Leu
305                 310                 315                 320

Ala Ser Ser Val Ser Phe His Ile Ala Glu Gly Val Arg Cys Asp Leu
                325                 330                 335

Arg Leu Asn Phe Ala Gly Val Pro Val Glu Lys Leu Asp Glu Cys Ala
            340                 345                 350

Pro Lys Phe Phe Asp Lys Thr Val Arg Glu His Leu Glu Glu Ala Asn
        355                 360                 365

Phe Asp Met Glu Arg Met Gly Tyr Leu Ile Asp Gln Thr Ile Leu Asn
    370                 375                 380

Glu Leu Val Lys Leu Glu Thr Asn Ala Pro Lys Asp Ile Met Ser His
385                 390                 395                 400

Ile Ile Gly His Gln Leu Phe Asp Asn Glu Asp Glu Glu Leu Phe Lys
                405                 410                 415

Lys Arg Thr Asn Glu Ile Asp Phe Leu Lys Lys Leu Lys Ser Glu Pro
            420                 425                 430

Ala Ser Tyr Trp Val Gln Leu Val Asn Lys Tyr Phe Thr Ala Pro Ser
        435                 440                 445
```

```
Ala Thr Val Ile Gly Val Pro Asn Glu Glu Leu Val Asp Lys Ile Ala
    450                 455                 460

Glu Glu Glu Glu Lys Arg Ile Ala Ala Gln Cys Glu Lys Leu Gly Lys
465                 470                 475                 480

Lys Gly Leu Glu Glu Ala Gly Lys Ser Leu Glu Ala Ala Ile Leu Glu
                485                 490                 495

Asn Thr Ala Asn His Pro Ser Ala Glu Leu Leu Asp Gln Leu Ile Val
            500                 505                 510

Lys Asp Leu Glu Ala Phe Asp Arg Phe Pro Val Gln Ser Leu Thr Ser
        515                 520                 525

Asn Ser Pro Ser Leu Thr Pro Gln Gln Ser Thr Phe Leu Ala Gln Phe
    530                 535                 540

Pro Phe His Ala Asn Leu His Asn Cys Pro Thr Lys Phe Val Glu Ile
545                 550                 555                 560

Phe Phe Leu Leu Asp Ser Ser Asn Leu Ser Ile Glu Asp Arg Ser Tyr
                565                 570                 575

Leu Phe Leu Tyr Thr Asp Leu Leu Phe Glu Ser Pro Ala Met Ile Asp
            580                 585                 590

Gly Val Leu Thr Ser Ala Asp Asp Val Ala Lys His Phe Thr Lys Asp
        595                 600                 605

Leu Ile Asp His Ser Ile Gln Val Gly Val Ser Gly Leu Tyr Asp Arg
    610                 615                 620

Phe Val Asn Leu Arg Ile Lys Val Gly Ala Asp Lys Tyr Pro Leu Leu
625                 630                 635                 640

Ala Lys Trp Ala Gln Ile Phe Thr Gln Gly Val Val Phe Asp Pro Ser
                645                 650                 655

Arg Ile His Gln Cys Ala Gln Lys Leu Ala Gly Glu Ala Arg Asp Arg
            660                 665                 670

Lys Arg Asp Gly Cys Thr Val Ala Ser Thr Ala Val Ala Ser Met Val
        675                 680                 685
```

```
Tyr Gly Lys Asn Thr Asn Cys Ile Leu Phe Asp Glu Leu Val Leu Glu
    690                 695                 700

Lys Leu His Glu Lys Ile Ser Lys Asp Val Met Lys Asn Pro Glu Ala
705                 710                 715                 720

Val Leu Glu Lys Leu Glu Gln Val Arg Ser Ala Leu Phe Ser Asn Gly
                725                 730                 735

Val Asn Ala His Phe Val Ala Asp Val Asp Ser Ile Asp Pro Lys Met
            740                 745                 750

Leu Ser Ser Asp Leu Trp Thr Trp Val Gln Ala Asp Pro Arg Phe Gly
        755                 760                 765

Pro Gly His Gln Phe Ser Ala Glu Ala Gly Glu Asn Val Ser Leu Glu
    770                 775                 780

Leu Gly Lys Glu Leu Leu Ile Gly Val Gly Gly Ser Glu Ser Ser Phe
785                 790                 795                 800

Ile Tyr Gln Thr Ser Phe Leu Asp Ala Asn Trp Asn Ser Glu Glu Leu
                805                 810                 815

Ile Pro Ala Met Ile Phe Gly Gln Tyr Leu Ser Gln Cys Glu Gly Pro
            820                 825                 830

Leu Trp Arg Ala Ile Arg Gly Asp Gly Leu Ala Tyr Gly Ala Asn Val
        835                 840                 845

Phe Val Lys Pro Asp Arg Lys Gln Ile Thr Leu Ser Leu Tyr Arg Cys
    850                 855                 860

Ala Gln Pro Ala Val Ala Tyr Glu Arg Thr Arg Asp Ile Ile Arg Lys
865                 870                 875                 880

Ile Val Glu Ser Gly Glu Ile Ser Lys Ala Glu Phe Glu Gly Ala Lys
                885                 890                 895

Arg Ser Thr Val Phe Glu Met Met Lys Arg Glu Gly Thr Val Ser Gly
            900                 905                 910

Ala Ala Lys Ile Ser Ile Leu Asn Asn Phe Arg Gln Thr Pro His Pro
        915                 920                 925
```

```
Phe Asn Ile Asp Leu Cys Arg Arg Ile Trp Asn Leu Thr Ser Glu Glu
    930                 935                 940

Met Val Lys Ile Gly Gly Pro Pro Leu Ala Arg Leu Phe Asp Glu Lys
945                 950                 955                 960

Cys Phe Val Arg Ser Ile Ala Val His Pro Ser Lys Leu Asn Glu Met
                965                 970                 975

Lys Lys Ala Phe Pro Gly Ser Ser Lys Ile Lys Ile Ser Asp Leu Gln
            980                 985                 990

Phe Ala Cys
        995
```

<210> 9
<211> 962
<212> PRT
<213> Escherichia coli

<400> 9

```
Met Pro Arg Ser Thr Trp Phe Lys Ala Leu Leu Leu Leu Val Ala Leu
1               5                   10                  15

Trp Ala Pro Leu Ser Gln Ala Glu Thr Gly Trp Gln Pro Ile Gln Glu
            20                  25                  30

Thr Ile Arg Lys Ser Asp Lys Asp Asn Arg Gln Tyr Gln Ala Ile Arg
        35                  40                  45

Leu Asp Asn Gly Met Val Val Leu Leu Val Ser Asp Pro Gln Ala Val
    50                  55                  60

Lys Ser Leu Ser Ala Leu Val Val Pro Val Gly Ser Leu Glu Asp Pro
65                  70                  75                  80

Glu Ala Tyr Gln Gly Leu Ala His Tyr Leu Glu His Met Ser Leu Met
                85                  90                  95

Gly Ser Lys Lys Tyr Pro Gln Ala Asp Ser Leu Ala Glu Tyr Leu Lys
            100                 105                 110

Met His Gly Gly Ser His Asn Ala Ser Thr Ala Pro Tyr Arg Thr Ala
        115                 120                 125
```

```
Phe Tyr Leu Glu Val Glu Asn Asp Ala Leu Pro Gly Ala Val Asp Arg
    130                 135                 140

Leu Ala Asp Ala Ile Ala Glu Pro Leu Leu Asp Lys Lys Tyr Ala Glu
145                 150                 155                 160

Arg Glu Arg Asn Ala Val Asn Ala Glu Leu Thr Met Ala Arg Thr Arg
                165                 170                 175

Asp Gly Met Arg Met Ala Gln Val Ser Ala Glu Thr Ile Asn Pro Ala
            180                 185                 190

His Pro Gly Ser Lys Phe Ser Gly Gly Asn Leu Glu Thr Leu Ser Asp
        195                 200                 205

Lys Pro Gly Asn Pro Val Gln Gln Ala Leu Lys Asp Phe His Glu Lys
    210                 215                 220

Tyr Tyr Ser Ala Asn Leu Met Lys Ala Val Ile Tyr Ser Asn Lys Pro
225                 230                 235                 240

Leu Pro Glu Leu Ala Lys Met Ala Ala Asp Thr Phe Gly Arg Val Pro
                245                 250                 255

Asn Lys Glu Ser Lys Lys Pro Glu Ile Thr Val Pro Val Val Thr Asp
            260                 265                 270

Ala Gln Lys Gly Ile Ile Ile His Tyr Val Pro Ala Leu Pro Arg Lys
        275                 280                 285

Val Leu Arg Val Glu Phe Arg Ile Asp Asn Asn Ser Ala Lys Phe Arg
    290                 295                 300

Ser Lys Thr Asp Glu Leu Ile Thr Tyr Leu Ile Gly Asn Arg Ser Pro
305                 310                 315                 320

Gly Thr Leu Ser Asp Trp Leu Gln Lys Gln Gly Leu Val Glu Gly Ile
                325                 330                 335

Ser Ala Asn Ser Asp Pro Ile Val Asn Gly Asn Ser Gly Val Leu Ala
            340                 345                 350

Ile Ser Ala Ser Leu Thr Asp Lys Gly Leu Ala Asn Arg Asp Gln Val
        355                 360                 365
```

```
Val Ala Ala Ile Phe Ser Tyr Leu Asn Leu Leu Arg Glu Lys Gly Ile
    370                 375                 380

Asp Lys Gln Tyr Phe Asp Glu Leu Ala Asn Val Leu Asp Ile Asp Phe
385                 390                 395                 400

Arg Tyr Pro Ser Ile Thr Arg Asp Met Asp Tyr Val Glu Trp Leu Ala
                405                 410                 415

Asp Thr Met Ile Arg Val Pro Val Glu His Thr Leu Asp Ala Val Asn
            420                 425                 430

Ile Ala Asp Arg Tyr Asp Ala Lys Ala Val Lys Glu Arg Leu Ala Met
        435                 440                 445

Met Thr Pro Gln Asn Ala Arg Ile Trp Tyr Ile Ser Pro Lys Glu Pro
    450                 455                 460

His Asn Lys Thr Ala Tyr Phe Val Asp Ala Pro Tyr Gln Val Asp Lys
465                 470                 475                 480

Ile Ser Ala Gln Thr Phe Ala Asp Trp Gln Lys Lys Ala Ala Asp Ile
                485                 490                 495

Ala Leu Ser Leu Pro Glu Leu Asn Pro Tyr Ile Pro Asp Asp Phe Ser
            500                 505                 510

Leu Ile Lys Ser Glu Lys Lys Tyr Asp His Pro Glu Leu Ile Val Asp
        515                 520                 525

Glu Ser Asn Leu Arg Val Val Tyr Ala Pro Ser Arg Tyr Phe Ala Ser
    530                 535                 540

Glu Pro Lys Ala Asp Val Ser Leu Ile Leu Arg Asn Pro Lys Ala Met
545                 550                 555                 560

Asp Ser Ala Arg Asn Gln Val Met Phe Ala Leu Asn Asp Tyr Leu Ala
                565                 570                 575

Gly Leu Ala Leu Asp Gln Leu Ser Asn Gln Ala Ser Val Gly Gly Ile
            580                 585                 590

Ser Phe Ser Thr Asn Ala Asn Asn Gly Leu Met Val Asn Ala Asn Gly
        595                 600                 605
```

```
Tyr Thr Gln Arg Leu Pro Gln Leu Phe Gln Ala Leu Leu Glu Gly Tyr
    610                 615                 620

Phe Ser Tyr Thr Ala Thr Glu Asp Gln Leu Glu Gln Ala Lys Ser Trp
625                 630                 635                 640

Tyr Asn Gln Met Met Asp Ser Ala Glu Lys Gly Lys Ala Phe Glu Gln
                645                 650                 655

Ala Ile Met Pro Ala Gln Met Leu Ser Gln Val Pro Tyr Phe Ser Arg
            660                 665                 670

Asp Glu Arg Arg Lys Ile Leu Pro Ser Ile Thr Leu Lys Glu Val Leu
        675                 680                 685

Ala Tyr Arg Asp Ala Leu Lys Ser Gly Ala Arg Pro Glu Phe Met Val
    690                 695                 700

Ile Gly Asn Met Thr Glu Ala Gln Ala Thr Thr Leu Ala Arg Asp Val
705                 710                 715                 720

Gln Lys Gln Leu Gly Ala Asp Gly Ser Glu Trp Cys Arg Asn Lys Asp
                725                 730                 735

Val Val Val Asp Lys Lys Gln Ser Val Ile Phe Glu Lys Ala Gly Asn
            740                 745                 750

Ser Thr Asp Ser Ala Leu Ala Ala Val Phe Val Pro Thr Gly Tyr Asp
        755                 760                 765

Glu Tyr Thr Ser Ser Ala Tyr Ser Ser Leu Leu Gly Gln Ile Val Gln
    770                 775                 780

Pro Trp Phe Tyr Asn Gln Leu Arg Thr Glu Glu Gln Leu Gly Tyr Ala
785                 790                 795                 800

Val Phe Ala Phe Pro Met Ser Val Gly Arg Gln Trp Gly Met Gly Phe
                805                 810                 815

Leu Leu Gln Ser Asn Asp Lys Gln Pro Ser Phe Leu Trp Glu Arg Tyr
            820                 825                 830

Lys Ala Phe Phe Pro Thr Ala Glu Ala Lys Leu Arg Ala Met Lys Pro
        835                 840                 845
```

```
Asp Glu Phe Ala Gln Ile Gln Gln Ala Val Ile Thr Gln Met Leu Gln
     850                 855                 860

Ala Pro Gln Thr Leu Gly Glu Glu Ala Ser Lys Leu Ser Lys Asp Phe
865                 870                 875                 880

Asp Arg Gly Asn Met Arg Phe Asp Ser Arg Asp Lys Ile Val Ala Gln
                885                 890                 895

Ile Lys Leu Leu Thr Pro Gln Lys Leu Ala Asp Phe Phe His Gln Ala
            900                 905                 910

Val Val Glu Pro Gln Gly Met Ala Ile Leu Ser Gln Ile Ser Gly Ser
        915                 920                 925

Gln Asn Gly Lys Ala Glu Tyr Val His Pro Glu Gly Trp Lys Val Trp
    930                 935                 940

Glu Asn Val Ser Ala Leu Gln Gln Thr Met Pro Leu Met Ser Glu Lys
945                 950                 955                 960

Asn Glu
```

<210> 10
<211> 1161
<212> PRT
<213> Homo sapiens

<400> 10

```
Met Leu Arg Arg Val Ala Val Ala Ala Val Phe Ala Thr Gly Arg Lys
1               5                   10                  15

Leu Arg Cys Glu Ala Gly Arg Asp Val Thr Ala Val Gly Arg Ile Glu
            20                  25                  30

Ala Arg Gly Leu Cys Glu Glu Ser Ala Lys Pro Phe Pro Thr Leu Thr
        35                  40                  45

Met Pro Gly Arg Asn Lys Ala Lys Ser Thr Cys Ser Cys Pro Asp Leu
    50                  55                  60

Gln Pro Asn Gly Gln Asp Leu Gly Glu Ser Gly Arg Val Ala Arg Leu
65                  70                  75                  80
```

```
Gly Ala Asp Glu Ser Glu Glu Glu Gly Arg Ser Leu Ser Asn Val Gly
                85                  90                  95

Asp Pro Glu Ile Ile Lys Ser Pro Ser Asp Pro Lys Gln Tyr Arg Tyr
            100                 105                 110

Ile Lys Leu Gln Asn Gly Leu Gln Ala Leu Leu Ile Ser Asp Leu Ser
        115                 120                 125

Asn Val Glu Gly Lys Thr Gly Asn Ala Thr Asp Glu Glu Glu Glu Glu
    130                 135                 140

Glu Glu Glu Glu Glu Glu Gly Glu Glu Glu Glu Glu Glu Glu Glu Asp
145                 150                 155                 160

Asp Asp Asp Asp Asp Asp Glu Asp Ser Gly Ala Glu Ile Gln Asp Asp
                165                 170                 175

Asp Glu Glu Gly Phe Asp Asp Glu Glu Glu Phe Asp Asp Asp Glu His
            180                 185                 190

Asp Asp Asp Asp Leu Asp Asn Glu Glu Asn Glu Leu Glu Glu Leu Glu
        195                 200                 205

Glu Arg Val Glu Ala Arg Lys Lys Thr Thr Glu Lys Gln Ser Ala Ala
    210                 215                 220

Ala Leu Cys Val Gly Val Gly Ser Phe Ala Asp Pro Asp Asp Leu Pro
225                 230                 235                 240

Gly Leu Ala His Phe Leu Glu His Met Val Phe Met Gly Ser Leu Lys
                245                 250                 255

Tyr Pro Asp Glu Asn Gly Phe Asp Ala Phe Leu Lys Lys His Gly Gly
            260                 265                 270

Ser Asp Asn Ala Ser Thr Asp Cys Glu Arg Thr Val Phe Gln Phe Asp
        275                 280                 285

Val Gln Arg Lys Tyr Phe Lys Glu Ala Leu Asp Arg Trp Ala Gln Phe
    290                 295                 300

Phe Ile His Pro Leu Met Ile Arg Asp Ala Ile Asp Arg Glu Val Glu
305                 310                 315                 320
```

```
Ala Val Asp Ser Glu Tyr Gln Leu Ala Arg Pro Ser Asp Ala Asn Arg
                325                 330                 335

Lys Glu Met Leu Phe Gly Ser Leu Ala Arg Pro Gly His Pro Met Gly
            340                 345                 350

Lys Phe Phe Trp Gly Asn Ala Glu Thr Leu Lys His Glu Pro Lys Lys
        355                 360                 365

Asn Asn Ile Asp Thr His Ala Arg Leu Arg Glu Phe Trp Met Arg Tyr
    370                 375                 380

Tyr Ser Ala His Tyr Met Thr Leu Val Val Gln Ser Lys Glu Thr Leu
385                 390                 395                 400

Asp Thr Leu Glu Lys Trp Val Thr Glu Ile Phe Ser Gln Ile Pro Asn
                405                 410                 415

Asn Gly Leu Pro Lys Pro Asn Phe Ser His Leu Thr Asp Pro Phe Asp
            420                 425                 430

Thr Pro Ala Phe Asn Lys Leu Tyr Arg Val Val Pro Ile Arg Lys Ile
        435                 440                 445

His Ala Leu Thr Ile Thr Trp Ala Leu Pro Pro Gln Gln Gln His Tyr
    450                 455                 460

Arg Val Lys Pro Leu His Tyr Ile Ser Trp Leu Val Gly His Glu Gly
465                 470                 475                 480

Lys Gly Ser Ile Leu Ser Tyr Leu Arg Lys Lys Cys Trp Ala Leu Ala
                485                 490                 495

Leu Phe Gly Gly Asn Gly Glu Thr Gly Phe Glu Gln Asn Ser Thr Tyr
            500                 505                 510

Ser Val Phe Ser Ile Ser Ile Thr Leu Thr Asp Glu Gly Tyr Glu His
        515                 520                 525

Phe Tyr Glu Val Ala His Thr Val Phe Gln Tyr Leu Lys Met Leu Gln
    530                 535                 540

Lys Leu Gly Pro Glu Lys Arg Val Phe Glu Glu Ile Gln Lys Ile Glu
545                 550                 555                 560
```

```
Asp Asn Glu Phe His Tyr Gln Glu Gln Thr Asp Pro Val Glu Tyr Val
                565                 570                 575

Glu Asn Met Cys Glu Asn Met Gln Leu Tyr Pro Arg Gln Asp Phe Leu
            580                 585                 590

Thr Gly Asp Gln Leu Leu Phe Glu Tyr Lys Pro Glu Val Ile Ala Glu
        595                 600                 605

Ala Leu Asn Gln Leu Val Pro Gln Lys Ala Asn Leu Val Leu Leu Ser
    610                 615                 620

Gly Ala Asn Glu Gly Arg Cys Asp Leu Lys Glu Lys Trp Phe Gly Thr
625                 630                 635                 640

Gln Tyr Ser Ile Glu Asp Ile Glu Asn Ser Trp Thr Glu Leu Trp Lys
                645                 650                 655

Ser Asn Phe Asp Leu Asn Ser Asp Leu His Leu Pro Ala Glu Asn Lys
            660                 665                 670

Tyr Ile Ala Thr Asp Phe Thr Leu Lys Ala Phe Asp Cys Pro Glu Thr
        675                 680                 685

Glu Tyr Pro Ala Lys Ile Val Asn Thr Pro Gln Gly Cys Leu Trp Tyr
    690                 695                 700

Lys Lys Asp Asn Lys Phe Lys Ile Pro Lys Ala Tyr Ile Arg Phe His
705                 710                 715                 720

Leu Ile Ser Pro Leu Ile Gln Lys Ser Ala Ala Asn Val Val Leu Phe
                725                 730                 735

Asp Ile Phe Val Asn Ile Leu Thr His Asn Leu Ala Glu Pro Ala Tyr
            740                 745                 750

Glu Ala Asp Val Ala Gln Leu Glu Tyr Lys Leu Val Ala Gly Glu His
        755                 760                 765

Gly Leu Ile Ile Arg Val Lys Gly Phe Asn His Lys Leu Pro Leu Leu
    770                 775                 780

Phe Gln Leu Ile Ile Asp Tyr Leu Thr Glu Phe Ser Ser Thr Pro Ala
785                 790                 795                 800
```

```
Val Phe Thr Met Ile Thr Glu Gln Leu Lys Lys Thr Tyr Phe Asn Ile
                805                 810                 815

Leu Ile Lys Pro Glu Thr Leu Ala Lys Asp Val Arg Leu Leu Ile Leu
            820                 825                 830

Glu Tyr Ser Arg Trp Ser Met Ile Asp Lys Tyr Arg Ala Leu Met Asp
        835                 840                 845

Gly Leu Ser Leu Glu Ser Leu Leu Asn Phe Val Lys Asp Phe Lys Ser
    850                 855                 860

Gln Leu Phe Val Glu Gly Leu Val Gln Gly Asn Val Thr Ser Thr Glu
865                 870                 875                 880

Ser Met Asp Phe Leu Arg Tyr Val Val Asp Lys Leu Asn Phe Val Pro
                885                 890                 895

Leu Glu Arg Glu Met Pro Val Gln Phe Gln Val Val Glu Leu Pro Ser
            900                 905                 910

Gly His His Leu Cys Lys Val Arg Ala Leu Asn Lys Gly Asp Ala Asn
        915                 920                 925

Ser Glu Val Thr Val Tyr Tyr Gln Ser Gly Thr Arg Ser Leu Arg Glu
    930                 935                 940

Tyr Thr Leu Met Glu Leu Leu Val Met His Met Glu Glu Pro Cys Phe
945                 950                 955                 960

Asp Phe Leu Arg Thr Lys Gln Thr Leu Gly Tyr His Val Tyr Pro Thr
                965                 970                 975

Cys Arg Asn Thr Ser Gly Ile Leu Gly Phe Ser Val Thr Val Gly Thr
            980                 985                 990

Gln Ala Thr Lys Tyr Asn Ser Glu  Thr Val Asp Lys Lys  Ile Glu Glu
        995                 1000                1005

Phe Leu  Ser Ser Phe Glu Glu  Lys Ile Glu Asn Leu  Thr Glu Asp
    1010                1015                1020

Ala Phe  Asn Thr Gln Val Thr  Ala Leu Ile Lys Leu  Lys Glu Cys
    1025                1030                1035
```

```
Glu Asp  Thr His Leu Gly Glu  Glu Val Asp Arg Asn  Trp Asn Glu
    1040                 1045                 1050

Val Val  Thr Gln Gln Tyr Leu  Phe Asp Arg Leu Ala  His Glu Ile
    1055                 1060                 1065

Glu Ala  Leu Lys Ser Phe Ser  Lys Ser Asp Leu Val  Ser Trp Phe
    1070                 1075                 1080

Lys Ala  His Arg Gly Pro Gly  Ser Lys Met Leu Ser  Val His Val
    1085                 1090                 1095

Val Gly  Tyr Gly Lys Tyr Glu  Leu Glu Glu Asp Gly  Ala Pro Val
    1100                 1105                 1110

Cys Glu  Asp Pro Asn Ser Arg  Glu Gly Met Gln Leu  Ile Tyr Leu
    1115                 1120                 1125

Pro Pro  Ser Pro Leu Leu Ala  Glu Ser Thr Thr Pro  Ile Thr Asp
    1130                 1135                 1140

Ile Arg  Ala Phe Thr Ala Thr  Leu Ser Leu Phe Pro  Tyr His Lys
    1145                 1150                 1155

Ile Val  Lys
    1160
```

<210> 11
<211> 1019
<212> PRT
<213> Homo sapiens

<400> 11

```
Met Arg Tyr Arg Leu Ala Trp Leu Leu His Pro Ala Leu Pro Ser Thr
1               5                  10                  15

Phe Arg Ser Val Leu Gly Ala Arg Leu Pro Pro Pro Glu Arg Leu Cys
            20                  25                  30

Gly Phe Gln Lys Lys Thr Tyr Ser Lys Met Asn Asn Pro Ala Ile Lys
        35                  40                  45

Arg Ile Gly Asn His Ile Thr Lys Ser Pro Glu Asp Lys Arg Glu Tyr
    50                  55                  60
```

```
Arg Gly Leu Glu Leu Ala Asn Gly Ile Lys Val Leu Leu Met Ser Asp
65                  70                  75                  80

Pro Thr Thr Asp Lys Ser Ser Ala Ala Leu Asp Val His Ile Gly Ser
                85                  90                  95

Leu Ser Asp Pro Pro Asn Ile Ala Gly Leu Ser His Phe Cys Glu His
            100                 105                 110

Met Leu Phe Leu Gly Thr Lys Lys Tyr Pro Lys Glu Asn Glu Tyr Ser
        115                 120                 125

Gln Phe Leu Ser Glu His Ala Gly Ser Ser Asn Ala Phe Thr Ser Gly
    130                 135                 140

Glu His Thr Asn Tyr Tyr Phe Asp Val Ser His Glu His Leu Glu Gly
145                 150                 155                 160

Ala Leu Asp Arg Phe Ala Gln Phe Phe Leu Cys Pro Leu Phe Asp Glu
                165                 170                 175

Ser Cys Lys Asp Arg Glu Val Asn Ala Val Asp Ser Glu His Glu Lys
            180                 185                 190

Asn Val Met Asn Asp Ala Trp Arg Leu Phe Gln Leu Glu Lys Ala Thr
        195                 200                 205

Gly Asn Pro Lys His Pro Phe Ser Lys Phe Gly Thr Gly Asn Lys Tyr
    210                 215                 220

Thr Leu Glu Thr Arg Pro Asn Gln Glu Gly Ile Asp Val Arg Gln Glu
225                 230                 235                 240

Leu Leu Lys Phe His Ser Ala Tyr Tyr Ser Ser Asn Leu Met Ala Val
                245                 250                 255

Cys Val Leu Gly Arg Glu Ser Leu Asp Asp Leu Thr Asn Leu Val Val
            260                 265                 270

Lys Leu Phe Ser Glu Val Glu Asn Lys Asn Val Pro Leu Pro Glu Phe
        275                 280                 285

Pro Glu His Pro Phe Gln Glu Glu His Leu Lys Gln Leu Tyr Lys Ile
    290                 295                 300
```

```
Val Pro Ile Lys Asp Ile Arg Asn Leu Tyr Val Thr Phe Pro Ile Pro
305                 310                 315                 320

Asp Leu Gln Lys Tyr Tyr Lys Ser Asn Pro Gly His Tyr Leu Gly His
                325                 330                 335

Leu Ile Gly His Glu Gly Pro Gly Ser Leu Leu Ser Glu Leu Lys Ser
            340                 345                 350

Lys Gly Trp Val Asn Thr Leu Val Gly Gly Gln Lys Glu Gly Ala Arg
        355                 360                 365

Gly Phe Met Phe Phe Ile Ile Asn Val Asp Leu Thr Glu Glu Gly Leu
    370                 375                 380

Leu His Val Glu Asp Ile Ile Leu His Met Phe Gln Tyr Ile Gln Lys
385                 390                 395                 400

Leu Arg Ala Glu Gly Pro Gln Glu Trp Val Phe Gln Glu Cys Lys Asp
                405                 410                 415

Leu Asn Ala Val Ala Phe Arg Phe Lys Asp Lys Glu Arg Pro Arg Gly
            420                 425                 430

Tyr Thr Ser Lys Ile Ala Gly Ile Leu His Tyr Tyr Pro Leu Glu Glu
        435                 440                 445

Val Leu Thr Ala Glu Tyr Leu Leu Glu Glu Phe Arg Pro Asp Leu Ile
    450                 455                 460

Glu Met Val Leu Asp Lys Leu Arg Pro Glu Asn Val Arg Val Ala Ile
465                 470                 475                 480

Val Ser Lys Ser Phe Glu Gly Lys Thr Asp Arg Thr Glu Glu Trp Tyr
                485                 490                 495

Gly Thr Gln Tyr Lys Gln Glu Ala Ile Pro Asp Glu Val Ile Lys Lys
            500                 505                 510

Trp Gln Asn Ala Asp Leu Asn Gly Lys Phe Lys Leu Pro Thr Lys Asn
        515                 520                 525

Glu Phe Ile Pro Thr Asn Phe Glu Ile Leu Pro Leu Glu Lys Glu Ala
    530                 535                 540
```

```
Thr Pro Tyr Pro Ala Leu Ile Lys Asp Thr Val Met Ser Lys Leu Trp
545                 550                 555                 560

Phe Lys Gln Asp Asp Lys Lys Lys Lys Pro Lys Ala Cys Leu Asn Phe
                565                 570                 575

Glu Phe Phe Ser Pro Phe Ala Tyr Val Asp Pro Leu His Cys Asn Met
            580                 585                 590

Ala Tyr Leu Tyr Leu Glu Leu Leu Lys Asp Ser Leu Asn Glu Tyr Ala
        595                 600                 605

Tyr Ala Ala Glu Leu Ala Gly Leu Ser Tyr Asp Leu Gln Asn Thr Ile
    610                 615                 620

Tyr Gly Met Tyr Leu Ser Val Lys Gly Tyr Asn Asp Lys Gln Pro Ile
625                 630                 635                 640

Leu Leu Lys Lys Ile Ile Glu Lys Met Ala Thr Phe Glu Ile Asp Glu
                645                 650                 655

Lys Arg Phe Glu Ile Ile Lys Glu Ala Tyr Met Arg Ser Leu Asn Asn
            660                 665                 670

Phe Arg Ala Glu Gln Pro His Gln His Ala Met Tyr Tyr Leu Arg Leu
        675                 680                 685

Leu Met Thr Glu Val Ala Trp Thr Lys Asp Glu Leu Lys Glu Ala Leu
    690                 695                 700

Asp Asp Val Thr Leu Pro Arg Leu Lys Ala Phe Ile Pro Gln Leu Leu
705                 710                 715                 720

Ser Arg Leu His Ile Glu Ala Leu Leu His Gly Asn Ile Thr Lys Gln
                725                 730                 735

Ala Ala Leu Gly Ile Met Gln Met Val Glu Asp Thr Leu Ile Glu His
            740                 745                 750

Ala His Thr Lys Pro Leu Leu Pro Ser Gln Leu Val Arg Tyr Arg Glu
        755                 760                 765

Val Gln Leu Pro Asp Arg Gly Trp Phe Val Tyr Gln Gln Arg Asn Glu
    770                 775                 780
```

```
Val His Asn Asn Cys Gly Ile Glu Ile Tyr Tyr Gln Thr Asp Met Gln
785                 790                 795                 800

Ser Thr Ser Glu Asn Met Phe Leu Glu Leu Phe Cys Gln Ile Ile Ser
                805                 810                 815

Glu Pro Cys Phe Asn Thr Leu Arg Thr Lys Glu Gln Leu Gly Tyr Ile
            820                 825                 830

Val Phe Ser Gly Pro Arg Arg Ala Asn Gly Ile Gln Ser Leu Arg Phe
        835                 840                 845

Ile Ile Gln Ser Glu Lys Pro Pro His Tyr Leu Glu Ser Arg Val Glu
    850                 855                 860

Ala Phe Leu Ile Thr Met Glu Lys Ser Ile Glu Asp Met Thr Glu Glu
865                 870                 875                 880

Ala Phe Gln Lys His Ile Gln Ala Leu Ala Ile Arg Arg Leu Asp Lys
                885                 890                 895

Pro Lys Lys Leu Ser Ala Glu Cys Ala Lys Tyr Trp Gly Glu Ile Ile
            900                 905                 910

Ser Gln Gln Tyr Asn Phe Asp Arg Asp Asn Thr Glu Val Ala Tyr Leu
        915                 920                 925

Lys Thr Leu Thr Lys Glu Asp Ile Ile Lys Phe Tyr Lys Glu Met Leu
    930                 935                 940

Ala Val Asp Ala Pro Arg Arg His Lys Val Ser Val His Val Leu Ala
945                 950                 955                 960

Arg Glu Met Asp Ser Cys Pro Val Val Gly Glu Phe Pro Cys Gln Asn
                965                 970                 975

Asp Ile Asn Leu Ser Gln Ala Pro Ala Leu Pro Gln Pro Glu Val Ile
            980                 985                 990

Gln Asn Met Thr Glu Phe Lys Arg  Gly Leu Pro Leu Phe  Pro Leu Val
        995                 1000                1005

Lys Pro  His Ile Asn Phe Met  Ala Ala Lys Leu
    1010                1015
```

<210> 12
<211> 1265
<212> PRT
<213> Arabidopsis thaliana

<400> 12

Met Ala Ser Ser Ser Ser Ser Ile Phe Thr Gly Val Lys Phe Ser Pro
1 5 10 15

Ile Leu Ala Pro Phe Asn Ser Gly Asp Ser Arg Arg Ser Arg Tyr Leu
20 25 30

Lys Asp Ser Arg Asn Lys Val Arg Phe Asn Pro Ser Ser Pro Arg Leu
35 40 45

Thr Pro His Arg Val Arg Val Glu Ala Pro Ser Leu Ile Pro Tyr Asn
50 55 60

Gly Leu Trp Ala Ala Gln Pro Asn Ser His Lys Gly Arg Leu Lys Arg
65 70 75 80

Asn Ile Val Ser Gly Lys Glu Ala Thr Gly Ile Ser Leu Ser Gln Gly
85 90 95

Arg Asn Phe Cys Leu Thr Cys Lys Arg Asn Gln Ala Gly Ile Arg Arg
100 105 110

Ala Leu Pro Ser Ala Phe Val Asp Arg Thr Ala Phe Ser Leu Ser Arg
115 120 125

Ser Ser Leu Thr Ser Ser Leu Arg Lys His Ser Gln Ile Val Asn Ala
130 135 140

Thr Leu Gly Pro Asp Glu Pro His Ala Ala Gly Thr Ala Trp Pro Asp
145 150 155 160

Gly Ile Val Ala Glu Arg Gln Asp Leu Asp Leu Leu Pro Pro Glu Ile
165 170 175

Asp Ser Ala Glu Leu Glu Ala Phe Leu Gly Cys Glu Leu Pro Ser His
180 185 190

Pro Lys Leu His Arg Gly Gln Leu Lys Asn Gly Leu Arg Tyr Leu Ile
195 200 205

Leu Pro Asn Lys Val Pro Pro Asn Arg Phe Glu Ala His Met Glu Val
210 215 220

His Val Gly Ser Ile Asp Glu Glu Glu Asp Glu Gln Gly Ile Ala His
225 230 235 240

Met Ile Glu His Val Ala Phe Leu Gly Ser Lys Lys Arg Glu Lys Leu
245 250 255

Leu Gly Thr Gly Ala Arg Ser Asn Ala Tyr Thr Asp Phe His His Thr
260 265 270

Val Phe His Ile His Ser Pro Thr His Thr Lys Asp Ser Glu Asp Asp
275 280 285

Leu Phe Pro Ser Val Leu Asp Ala Leu Asn Glu Ile Ala Phe His Pro
290 295 300

Lys Phe Leu Ser Ser Arg Val Glu Lys Glu Arg Arg Ala Ile Leu Ser
305 310 315 320

Glu Leu Gln Met Met Asn Thr Ile Glu Tyr Arg Val Asp Cys Gln Leu
325 330 335

Leu Gln His Leu His Ser Glu Asn Lys Leu Gly Arg Arg Phe Pro Ile
340 345 350

Gly Leu Glu Glu Gln Ile Lys Lys Trp Asp Val Asp Lys Ile Arg Lys
355 360 365

Phe His Glu Arg Trp Tyr Phe Pro Ala Asn Ala Thr Leu Tyr Ile Val
370 375 380

Gly Asp Ile Asp Asn Ile Pro Arg Ile Val His Asn Ile Glu Ala Val
385 390 395 400

Phe Gly Lys Asn Gly Leu Asp Asn Glu Ser Thr Pro Ser Ser Pro Ser
405 410 415

Pro Gly Ala Phe Gly Ala Met Ala Asn Phe Leu Val Pro Lys Leu Pro
420 425 430

Ala Gly Leu Gly Gly Thr Phe Ser Asn Glu Lys Thr Asn Thr Ala Asp
435 440 445

```
Gln Ser Lys Met Ile Lys Arg Glu Arg His Ala Ile Arg Pro Pro Val
    450                 455                 460

Glu His Asn Trp Ser Leu Pro Gly Thr Ser Val Asp Leu Lys Pro Pro
465                 470                 475                 480

Gln Ile Phe Lys His Glu Leu Leu Gln Asn Phe Ala Ile Asn Met Phe
                485                 490                 495

Cys Lys Ile Pro Val Ser Lys Val Gln Thr Phe Gly Asp Leu Arg Asn
            500                 505                 510

Val Leu Met Lys Arg Ile Phe Leu Ser Ala Leu His Phe Arg Ile Asn
        515                 520                 525

Thr Arg Tyr Lys Ser Ser Asn Pro Pro Phe Thr Ser Val Glu Leu Asp
    530                 535                 540

His Ser Asp Ser Gly Arg Glu Gly Cys Thr Val Thr Thr Leu Thr Val
545                 550                 555                 560

Thr Ala Glu Pro Gln Asn Trp Gln Asn Ala Val Lys Val Ala Val Gln
                565                 570                 575

Glu Val Arg Arg Leu Lys Glu Phe Gly Val Thr Arg Gly Glu Leu Thr
            580                 585                 590

Arg Tyr Met Asp Ala Leu Leu Lys Asp Ser Glu His Leu Ala Ala Met
        595                 600                 605

Ile Asp Asn Val Ser Ser Val Asp Asn Leu Asp Phe Ile Met Glu Ser
    610                 615                 620

Asp Ala Leu Ser His Thr Val Met Asp Gln Thr Gln Gly His Glu Thr
625                 630                 635                 640

Leu Val Ala Val Ala Gly Thr Val Thr Leu Glu Glu Val Asn Thr Val
                645                 650                 655

Gly Ala Lys Val Leu Glu Phe Ile Ser Asp Phe Gly Arg Pro Thr Ala
            660                 665                 670

Leu Leu Pro Ala Ala Ile Val Ala Cys Val Pro Thr Lys Val His Val
        675                 680                 685
```

```
Asp Gly Val Gly Glu Ser Asp Phe Asn Ile Ser Pro Asp Glu Ile Ile
    690                 695                 700

Glu Ser Val Lys Ser Gly Leu Leu Ala Pro Ile Glu Ala Glu Pro Glu
705                 710                 715                 720

Leu Glu Val Pro Lys Glu Leu Ile Ser Gln Ser Gln Leu Lys Glu Leu
                725                 730                 735

Thr Leu Gln Arg Asn Pro Cys Phe Val Pro Ile Pro Gly Ser Gly Leu
            740                 745                 750

Thr Lys Leu His Asp Lys Glu Thr Gly Ile Thr Gln Leu Arg Leu Ser
        755                 760                 765

Asn Gly Ile Ala Val Asn Tyr Lys Lys Ser Thr Thr Glu Ser Arg Ala
    770                 775                 780

Gly Val Met Arg Leu Ile Val Gly Gly Gly Arg Ala Ala Glu Thr Ser
785                 790                 795                 800

Asp Ser Lys Gly Ala Val Val Val Gly Val Arg Thr Leu Ser Glu Gly
                805                 810                 815

Gly Arg Val Gly Asp Phe Ser Arg Glu Gln Val Glu Leu Phe Cys Val
            820                 825                 830

Asn His Leu Ile Asn Cys Ser Leu Glu Ser Thr Glu Glu Phe Ile Ala
        835                 840                 845

Met Glu Phe Arg Phe Thr Leu Arg Asp Asn Gly Met Gln Ala Ala Phe
    850                 855                 860

Gln Leu Leu His Met Val Leu Glu Arg Ser Val Trp Leu Glu Asp Ala
865                 870                 875                 880

Phe Asp Arg Ala Arg Gln Leu Tyr Leu Ser Tyr Phe Arg Ser Ile Pro
                885                 890                 895

Lys Ser Leu Glu Arg Ala Thr Ala His Lys Leu Met Ile Ala Met Leu
            900                 905                 910

Asn Gly Asp Glu Arg Phe Val Glu Pro Thr Pro Lys Ser Leu Gln Ser
        915                 920                 925
```

```
Leu Asn Leu Glu Ser Val Lys Asp Ala Val Met Ser His Phe Val Gly
    930                 935                 940

Asp Asn Met Glu Val Ser Ile Val Gly Asp Phe Ser Glu Glu Glu Ile
945                 950                 955                 960

Glu Arg Cys Ile Leu Asp Tyr Leu Gly Thr Val Lys Ala Ser His Asp
                965                 970                 975

Ser Ala Lys Pro Pro Gly Ser Glu Pro Ile Leu Phe Arg Gln Pro Thr
            980                 985                 990

Ala Gly Leu Gln Phe Gln Gln Val  Phe Leu Lys Asp Thr  Asp Glu Arg
        995                 1000                1005

Ala Cys  Ala Tyr Ile Ala Gly  Pro Ala Pro Asn Arg  Trp Gly Phe
    1010                1015                1020

Thr Val  Asp Gly Asp Asp Leu  Phe Gln Ser Val Ser  Lys Leu Pro
    1025                1030                1035

Val Ala  His Asp Gly Leu Leu  Lys Ser Glu Glu Gln  Leu Leu Glu
    1040                1045                1050

Gly Gly  Asp Arg Glu Leu Gln  Lys Lys Leu Arg Ala  His Pro Leu
    1055                1060                1065

Phe Phe  Gly Val Thr Met Gly  Leu Leu Ala Glu Ile  Ile Asn Ser
    1070                1075                1080

Arg Leu  Phe Thr Thr Val Arg  Asp Ser Leu Gly Leu  Thr Tyr Asp
    1085                1090                1095

Val Ser  Phe Glu Leu Asn Leu  Phe Asp Arg Leu Lys  Leu Gly Trp
    1100                1105                1110

Tyr Val  Ile Ser Val Thr Ser  Thr Pro Gly Lys Val  Tyr Lys Ala
    1115                1120                1125

Val Asp  Ala Cys Lys Asn Val  Leu Arg Gly Leu His  Ser Asn Gln
    1130                1135                1140

Ile Ala  Pro Arg Glu Leu Asp  Arg Ala Lys Arg Thr  Leu Leu Met
    1145                1150                1155
```

```
Arg His  Glu Ala Glu Leu Lys  Ser Asn Ala Tyr Trp  Leu Asn Leu
    1160                 1165                 1170

Leu Ala  His Leu Gln Ala Ser  Ser Val Gln Arg Lys  Glu Leu Ser
    1175                 1180                 1185

Cys Ile  Lys Glu Leu Val Ser  Leu Tyr Glu Ala Ala  Ser Ile Glu
    1190                 1195                 1200

Asp Ile  Tyr Leu Ala Tyr Asn  Gln Leu Arg Val Asp  Glu Asp Ser
    1205                 1210                 1215

Leu Tyr  Ser Cys Ile Gly Ile  Ala Gly Ala Gln Ala  Gly Glu Glu
    1220                 1225                 1230

Ile Thr  Val Leu Ser Glu Glu  Glu Glu Pro Glu Asp  Val Phe Ser
    1235                 1240                 1245

Gly Val  Val Pro Val Gly Arg  Gly Ser Ser Met Thr  Thr Arg Pro
    1250                 1255                 1260

Thr Thr
    1265
```

<210> 13
<211> 534
<212> PRT
<213> Homo sapiens

<400> 13

```
Met Arg Pro Asp Asp Lys Tyr His Glu Lys Gln Ala Gln Val Glu Ala
1               5                  10                  15

Thr Lys Leu Lys Gln Lys Val Glu Ala Leu Ser Pro Gly Asp Arg Gln
            20                  25                  30

Gln Ile Tyr Glu Lys Gly Leu Glu Leu Arg Ser Gln Gln Ser Lys Pro
        35                  40                  45

Gln Asp Ala Ser Cys Leu Pro Ala Leu Lys Val Ser Asp Ile Glu Pro
    50                  55                  60

Thr Ile Pro Val Thr Glu Leu Asp Val Val Leu Thr Ala Gly Asp Ile
65                  70                  75                  80
```

```
Pro Val Gln Tyr Cys Ala Gln Pro Thr Asn Gly Met Val Tyr Phe Arg
                85                  90                  95

Ala Phe Ser Ser Leu Asn Thr Leu Pro Glu Glu Leu Arg Pro Tyr Val
            100                 105                 110

Pro Leu Phe Cys Ser Val Leu Thr Lys Leu Gly Cys Gly Leu Leu Asp
        115                 120                 125

Tyr Arg Glu Gln Ala Gln Gln Ile Glu Leu Lys Thr Gly Gly Met Ser
    130                 135                 140

Ala Ser Pro His Val Leu Pro Asp Asp Ser His Met Asp Thr Tyr Glu
145                 150                 155                 160

Gln Gly Val Leu Phe Ser Ser Leu Cys Leu Asp Arg Asn Leu Pro Asp
                165                 170                 175

Met Met Gln Leu Trp Ser Glu Ile Phe Asn Asn Pro Cys Phe Glu Glu
            180                 185                 190

Glu Glu His Phe Lys Val Leu Val Lys Met Thr Ala Gln Glu Leu Ala
        195                 200                 205

Asn Gly Ile Pro Asp Ser Gly His Leu Tyr Ala Ser Ile Arg Ala Gly
    210                 215                 220

Arg Thr Leu Thr Pro Ala Gly Asp Leu Gln Glu Thr Phe Ser Gly Met
225                 230                 235                 240

Asp Gln Val Arg Leu Met Lys Arg Ile Ala Glu Met Thr Asp Ile Lys
                245                 250                 255

Pro Ile Leu Arg Lys Leu Pro Arg Ile Lys Lys His Leu Leu Asn Gly
            260                 265                 270

Asp Asn Met Arg Cys Ser Val Asn Ala Thr Pro Gln Gln Met Pro Gln
        275                 280                 285

Thr Glu Lys Ala Val Glu Asp Phe Leu Arg Ser Ile Gly Arg Ser Lys
    290                 295                 300

Lys Glu Arg Arg Pro Val Arg Pro His Thr Val Glu Lys Pro Val Pro
305                 310                 315                 320
```

```
Ser Ser Ser Gly Gly Asp Ala His Val Pro His Gly Ser Gln Val Ile
                325                 330                 335

Arg Lys Leu Val Met Glu Pro Thr Phe Lys Pro Trp Gln Met Lys Thr
            340                 345                 350

His Phe Leu Met Pro Phe Pro Val Asn Tyr Val Gly Glu Cys Ile Arg
        355                 360                 365

Thr Val Pro Tyr Thr Asp Pro Asp His Ala Ser Leu Lys Ile Leu Ala
    370                 375                 380

Arg Leu Met Thr Ala Lys Phe Leu His Thr Glu Ile Arg Glu Lys Gly
385                 390                 395                 400

Gly Ala Tyr Gly Gly Gly Ala Lys Leu Ser His Asn Gly Ile Phe Thr
                405                 410                 415

Leu Tyr Ser Tyr Arg Asp Pro Asn Thr Ile Glu Thr Leu Gln Ser Phe
            420                 425                 430

Gly Lys Ala Val Asp Trp Ala Lys Ser Gly Lys Phe Thr Gln Gln Asp
        435                 440                 445

Ile Asp Glu Ala Lys Leu Ser Val Phe Ser Thr Val Asp Ala Pro Val
    450                 455                 460

Ala Pro Ser Asp Lys Gly Met Asp His Phe Leu Tyr Gly Leu Ser Asp
465                 470                 475                 480

Glu Met Lys Gln Ala His Arg Glu Gln Leu Phe Ala Val Ser His Asp
                485                 490                 495

Lys Leu Leu Ala Val Ser Asp Arg Tyr Leu Gly Thr Gly Lys Ser Thr
            500                 505                 510

His Gly Leu Ala Ile Leu Gly Pro Glu Asn Pro Lys Ile Ala Lys Asp
        515                 520                 525

Pro Ser Trp Ile Ile Arg
    530
```

<210> 14
<211> 409
<212> PRT
<213> Bacillus subtilis

<400> 14

```
Met Ile Lys Arg Tyr Thr Cys Pro Asn Gly Val Arg Ile Val Leu Glu
1               5                  10                  15

Asn Asn Pro Thr Val Arg Ser Val Ala Ile Gly Val Trp Ile Gly Thr
            20                  25                  30

Gly Ser Arg His Glu Thr Pro Glu Ile Asn Gly Ile Ser His Phe Leu
        35                  40                  45

Glu His Met Phe Phe Lys Gly Thr Ser Thr Lys Ser Ala Arg Glu Ile
    50                  55                  60

Ala Glu Ser Phe Asp Arg Ile Gly Gly Gln Val Asn Ala Phe Thr Ser
65                  70                  75                  80

Lys Glu Tyr Thr Cys Tyr Tyr Ala Lys Val Leu Asp Glu His Ala Asn
                85                  90                  95

Tyr Ala Leu Asp Val Leu Ala Asp Met Phe Phe His Ser Thr Phe Asp
            100                 105                 110

Glu Asn Glu Leu Lys Lys Glu Lys Asn Val Val Tyr Glu Glu Ile Lys
        115                 120                 125

Met Tyr Glu Asp Ala Pro Asp Asp Ile Val His Asp Leu Leu Ser Lys
    130                 135                 140

Ala Thr Tyr Gly Asn His Ser Leu Gly Tyr Pro Ile Leu Gly Thr Glu
145                 150                 155                 160

Glu Thr Leu Ala Ser Phe Asn Gly Asp Ser Leu Arg Gln Tyr Met His
                165                 170                 175

Asp Tyr Tyr Thr Pro Asp Arg Val Val Ile Ser Val Ala Gly Asn Ile
            180                 185                 190

Ser Asp Ser Phe Ile Lys Asp Val Glu Lys Trp Phe Gly Ser Tyr Glu
        195                 200                 205
```

```
Ala Lys Gly Lys Ala Thr Gly Leu Glu Lys Pro Glu Phe His Thr Glu
    210                 215                 220

Lys Leu Thr Arg Lys Lys Glu Thr Glu Gln Ala His Leu Cys Leu Gly
225                 230                 235                 240

Phe Lys Gly Leu Glu Val Gly His Glu Arg Ile Tyr Asp Leu Ile Val
                245                 250                 255

Leu Asn Asn Val Leu Gly Gly Ser Met Ser Ser Arg Leu Phe Gln Asp
            260                 265                 270

Val Arg Glu Asp Lys Gly Leu Ala Tyr Ser Val Tyr Ser Tyr His Ser
        275                 280                 285

Ser Tyr Glu Asp Ser Gly Met Leu Thr Ile Tyr Gly Gly Thr Gly Ala
    290                 295                 300

Asn Gln Leu Gln Gln Leu Ser Glu Thr Ile Gln Glu Thr Leu Ala Thr
305                 310                 315                 320

Leu Lys Arg Asp Gly Ile Thr Ser Lys Glu Leu Glu Asn Ser Lys Glu
                325                 330                 335

Gln Met Lys Gly Ser Leu Met Leu Ser Leu Glu Ser Thr Asn Ser Lys
            340                 345                 350

Met Ser Arg Asn Gly Lys Asn Glu Leu Leu Leu Gly Lys His Lys Thr
        355                 360                 365

Leu Asp Glu Ile Ile Asn Glu Leu Asn Ala Val Asn Leu Glu Arg Val
    370                 375                 380

Asn Gly Leu Ala Arg Gln Leu Phe Thr Glu Asp Tyr Ala Leu Ala Leu
385                 390                 395                 400

Ile Ser Pro Ser Gly Asn Met Pro Ser
                405
```

<210> 15
<211> 438
<212> PRT
<213> Mycobacterium tuberculosis

<400> 15

```
Met Pro Arg Arg Ser Pro Ala Asp Pro Ala Ala Ala Leu Ala Pro Arg
1               5                  10                  15

Arg Thr Thr Leu Pro Gly Gly Leu Arg Val Val Thr Glu Phe Leu Pro
            20                  25                  30

Ala Val His Ser Ala Ser Val Gly Val Trp Val Gly Val Gly Ser Arg
        35                  40                  45

Asp Glu Gly Ala Thr Val Ala Gly Ala Ala His Phe Leu Glu His Leu
    50                  55                  60

Leu Phe Lys Ser Thr Pro Thr Arg Ser Ala Val Asp Ile Ala Gln Ala
65                  70                  75                  80

Met Asp Ala Val Gly Gly Glu Leu Asn Ala Phe Thr Ala Lys Glu His
                85                  90                  95

Thr Cys Tyr Tyr Ala His Val Leu Gly Ser Asp Leu Pro Leu Ala Val
            100                 105                 110

Asp Leu Val Ala Asp Val Val Leu Asn Gly Arg Cys Ala Ala Asp Asp
        115                 120                 125

Val Glu Val Glu Arg Asp Val Val Leu Glu Glu Ile Ala Met Arg Asp
    130                 135                 140

Asp Asp Pro Glu Asp Ala Leu Ala Asp Met Phe Leu Ala Ala Leu Phe
145                 150                 155                 160

Gly Asp His Pro Val Gly Arg Pro Val Ile Gly Ser Ala Gln Ser Val
                165                 170                 175

Ser Val Met Thr Arg Ala Gln Leu Gln Ser Phe His Leu Arg Arg Tyr
            180                 185                 190

Thr Pro Glu Arg Met Val Val Ala Ala Ala Gly Asn Val Asp His Asp
        195                 200                 205

Gly Leu Val Ala Leu Val Arg Glu His Phe Gly Ser Arg Leu Val Arg
    210                 215                 220

Gly Arg Arg Pro Val Ala Pro Arg Lys Gly Thr Gly Arg Val Asn Gly
225                 230                 235                 240
```

Ser Pro Arg Leu Thr Leu Val Ser Arg Asp Ala Glu Gln Thr His Val
245 250 255

Ser Leu Gly Ile Arg Thr Pro Gly Arg Gly Trp Glu His Arg Trp Ala
260 265 270

Leu Ser Val Leu His Thr Ala Leu Gly Gly Gly Leu Ser Ser Arg Leu
275 280 285

Phe Gln Glu Val Arg Glu Thr Arg Gly Leu Ala Tyr Ser Val Tyr Ser
290 295 300

Ala Leu Asp Leu Phe Ala Asp Ser Gly Ala Leu Ser Val Tyr Ala Ala
305 310 315 320

Cys Leu Pro Glu Arg Phe Ala Asp Val Met Arg Val Thr Ala Asp Val
325 330 335

Leu Glu Ser Val Ala Arg Asp Gly Ile Thr Glu Ala Glu Cys Gly Ile
340 345 350

Ala Lys Gly Ser Leu Arg Gly Gly Leu Val Leu Gly Leu Glu Asp Ser
355 360 365

Ser Ser Arg Met Ser Arg Leu Gly Arg Ser Glu Leu Asn Tyr Gly Lys
370 375 380

His Arg Ser Ile Glu His Thr Leu Arg Gln Ile Glu Gln Val Thr Val
385 390 395 400

Glu Glu Val Asn Ala Val Ala Arg His Leu Leu Ser Arg Arg Tyr Gly
405 410 415

Ala Ala Val Leu Gly Pro His Gly Ser Lys Arg Ser Leu Pro Gln Gln
420 425 430

Leu Arg Ala Met Val Gly
435

<210> 16
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 16
aatagaagct tgtcgactga tctatccaaa actg 34

<210> 17
<211> 66
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 17
aaaagagctc ggccagatct tctagaggat ccaagaattc tgttttatat ttgttgtaaa 60

aagtag 66

<210> 18
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 18
ttttgaattc caagatctcc catgtctcta ctggtgg 37

<210> 19
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 19
ccccgagctc gtcgaccctt ctcgaaagct ttaacgaacg c 41

<210> 20
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 20
ttttgaattc aaagaatgag atttccttca atttttactg cag 43

<210> 21
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 21
tttttctaga ctaggagggg tactcatact cctcggc 37

<210> 22
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 22
cgaatgtcca tcgttgcgaa cctgcagaac ctg 33

<210> 23
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 23
caggttctgc aggttcctaa cgatggacat tcg 33

<210> 24
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 24
cgaatgtcca tcgttaggaa cctgcagaac ctg 33

<210> 25
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 25
caggttctgc aggttcctaa cgatggacat tcg 33

<210> 26
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 26
tcgcagagaa cggatggc 18

<210> 27
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 27
ttttgggccc ttcatggtga tacggtatct cttggc 36

<210> 28
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 28
ttttctcgag aaggtggaac atactgccct gggatgg 37

<210> 29
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 29
ttttgagctc gtttaggaaa cgtccttggc ggagatgc 38

<210> 30
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 30
tttttctaga cactgcgaat ccatggtata aaccaaaacc 40

<210> 31
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 31
gtcgttgttc atggacatac ctcc 24

<210> 32
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 32
tacaaatgtt cttctgccat ttctgg 26

<210> 33
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 33
ggttcatatg cgccggagct cctcgacagc ag 32

<210> 34
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 34
ggttcctagg atccgcaagt ttgattccat tgcggtg 37

<210> 35
<211> 70
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 35
ttaaagagta ccttggctat agaataccgt agagataaag acctgaatag agattgtact 60

gagagtgcac 70

<210> 36
<211> 70
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 36
aggtattata actatttttc tgtatttttt atatattttt atttgccaag ctgtgcggta 60

tttcacaccg 70

<210> 37
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 37
ctttggttaa agagtacctt ggc 23

<210> 38
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 38
tactacgaaa agcgtgtgcg agg 23

<210> 39
<211> 71
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 39
tagaaggcta ctcaaaagaa taaagttact ataaaatata ctgcggtata tagattgtac 60

tgagagtgca c 71

<210> 40
<211> 70
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 40
gatcggcaag aaactttgaa gcagtatatt tacaggatta aattatatat ctgtgcggta 60

tttcacaccg 70

<210> 41
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 41
cggaggggct ctatgataaa gg 22

<210> 42
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 42
gagtaactag ggcttctctt ccc 23

<210> 43
<211> 85
<212> PRT
<213> Homo sapiens

<400> 43

Ser Gly Leu Gln Arg Ala Glu Glu Ala Pro Arg Arg Gln Leu Arg Val
1 5 10 15

Ser Gln Arg Thr Asp Gly Glu Ser Arg Ala His Leu Gly Ala Leu Leu
20 25 30

Ala Arg Tyr Ile Gln Gln Ala Arg Lys Ala Pro Ser Gly Arg Met Ser
35 40 45

Ile Val Lys Asn Leu Gln Asn Leu Asp Pro Ser His Arg Ile Ser Asp
50 55 60

Arg Asp Tyr Met Gly Trp Met Asp Phe Gly Arg Arg Ser Ala Glu Glu
65 70 75 80

Tyr Glu Tyr Pro Ser
85

<210> 44
<211> 22
<212> PRT
<213> Homo sapiens

<400> 44

Gln Leu Arg Val Ser Gln Arg Thr Asp Gly Glu Ser Arg Ala His Leu
1 5 10 15

Gly Ala Leu Leu Ala Arg
20

<210> 45
<211> 19
<212> PRT
<213> Homo sapiens

<400> 45

Val Ser Gln Arg Thr Asp Gly Glu Ser Arg Ala His Leu Gly Ala Leu
1 5 10 15

Leu Ala Arg

<210> 46
<211> 51
<212> PRT
<213> Homo sapiens

<400> 46

Tyr Ile Gln Gln Ala Arg Lys Ala Pro Ser Gly Arg Met Ser Ile Val
1 5 10 15

Lys Asn Leu Gln Asn Leu Asp Pro Ser His Arg Ile Ser Asp Arg Asp
20 25 30

Tyr Met Gly Trp Met Asp Phe Gly Arg Arg Ser Ala Glu Glu Tyr Glu
35 40 45

Tyr Pro Ser
50

<210> 47
<211> 17
<212> PRT
<213> Homo sapiens

<400> 47

Tyr Ile Gln Gln Ala Arg Lys Ala Pro Ser Gly Arg Met Ser Ile Val
1 5 10 15

Lys

<210> 48
<211> 16
<212> PRT
<213> Homo sapiens

<400> 48

Tyr Ile Gln Gln Ala Arg Lys Ala Pro Ser Gly Arg Met Ser Ile Val
1 5 10 15

<210> 49
<211> 13
<212> PRT
<213> Homo sapiens

<400> 49

Asn Leu Gln Asn Leu Asp Pro Ser His Arg Ile Ser Asp
1 5 10

<210> 50
<211> 23
<212> PRT
<213> Homo sapiens

<400> 50

Asn Leu Gln Asn Leu Asp Pro Ser His Arg Ile Ser Asp Arg Asp Tyr
1 5 10 15

Met Gly Trp Met Asp Phe Gly
20

<210> 51
<211> 34
<212> PRT
<213> Homo sapiens

<400> 51

Asn Leu Gln Asn Leu Asp Pro Ser His Arg Ile Ser Asp Arg Asp Tyr
1 5 10 15

Met Gly Trp Met Asp Phe Gly Arg Arg Ser Ala Glu Glu Tyr Glu Tyr
20 25 30

Pro Ser

<210> 53
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> Portion of fusion protein

<400> 53

Lys Arg Glu Ala Glu Ala Ser Gly Leu Gln Arg Ala
1 5 10

<210> 54
<211> 9
<212> PRT
<213> Homo sapiens

<400> 54

Arg Met Ser Ile Val Lys Asn Leu Gln
1 5

<210> 55
<211> 13
<212> PRT
<213> Homo sapiens

<400> 55

Asp Arg Asp Tyr Met Gly Trp Met Asp Phe Gly Arg Arg
1 5 10

This Page is inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☑ FADED TEXT OR DRAWING
- ☑ BLURED OR ILLEGIBLE TEXT OR DRAWING
- ☑ SKEWED/SLANTED IMAGES
- ☐ COLORED OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☑ REPERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning documents *will not* correct images problems checked, please do not report the problems to the IFW Image Problem Mailbox

Modtaget PVS
20 SEP. 2002

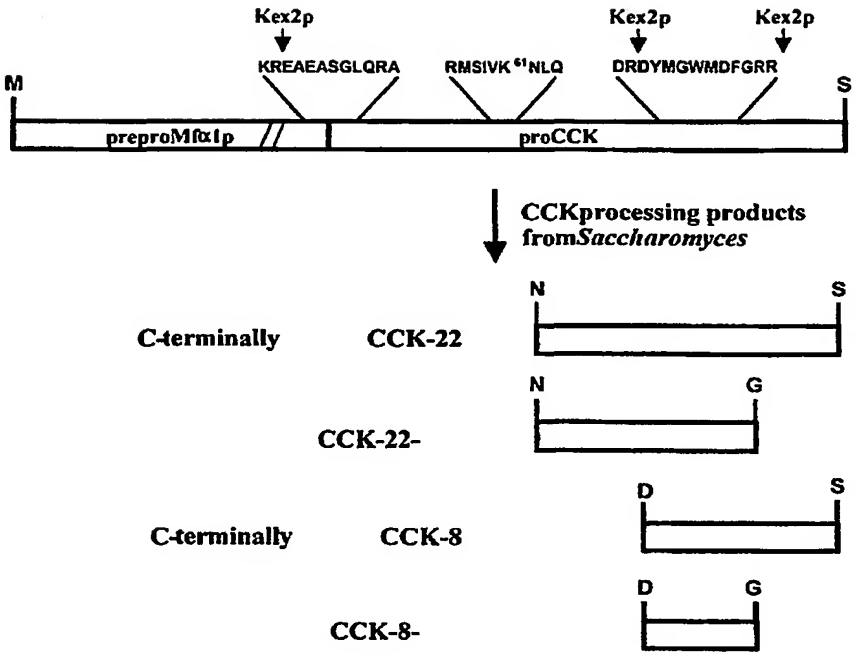

Figure 1

Modtaget PVS
20 SEP. 2002

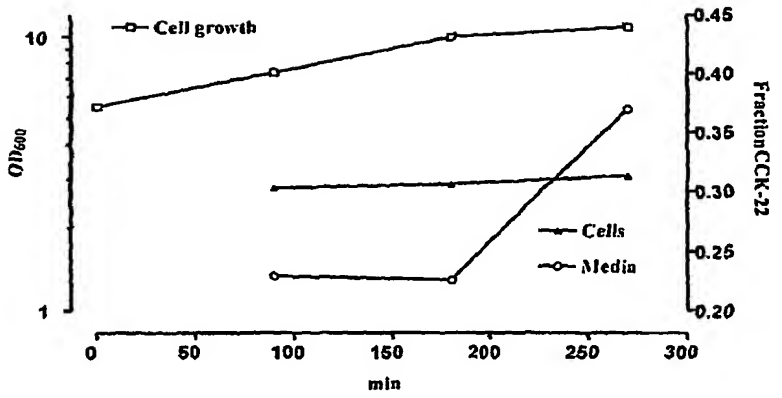

Figure 2

Modtaget PVS
20 SEP. 2002

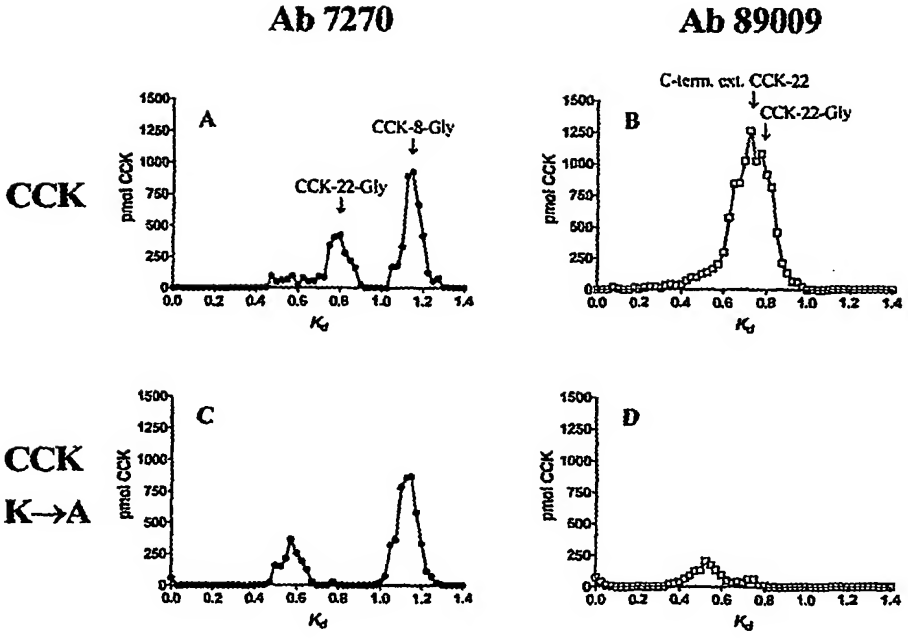

Figure 3

Modtaget PVS
20 SEP. 2002

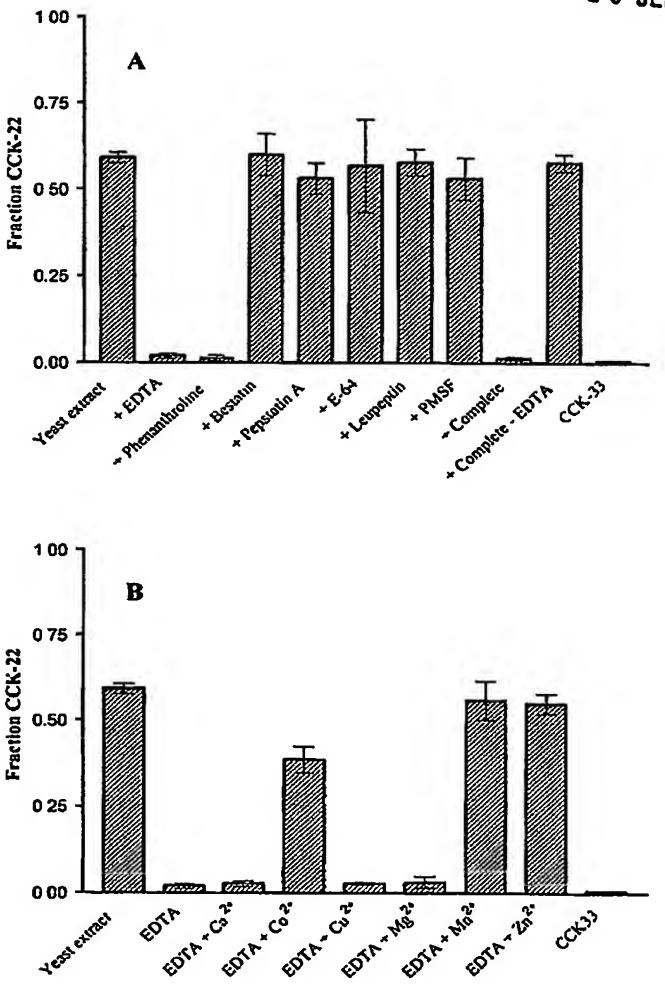

Figure 4

Modtaget PVS
20 SEP. 2002

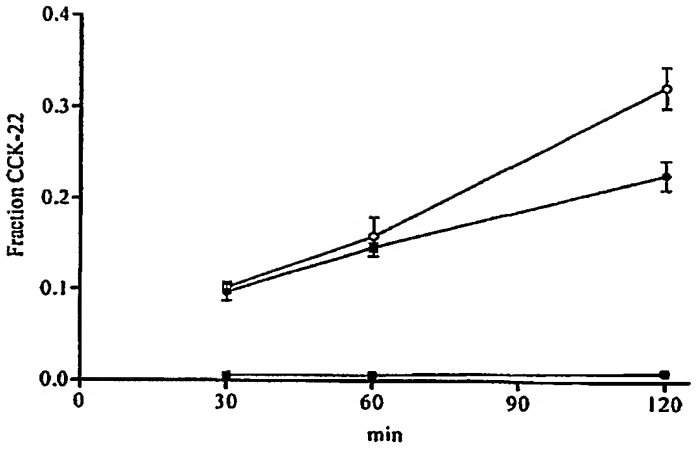

Figure 5

Modtaget PVS
20 SEP. 2002

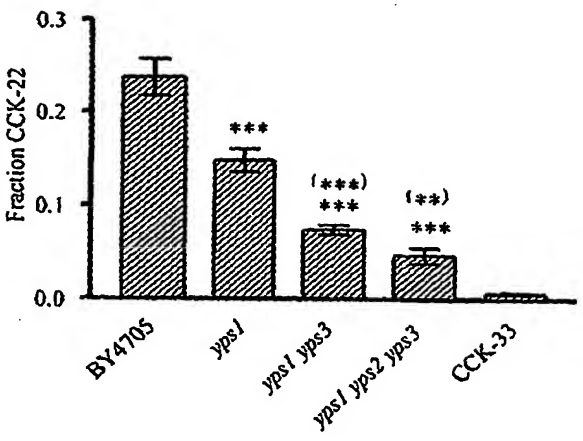

Figure 6

Modtaget PVS
20 SEP. 2002

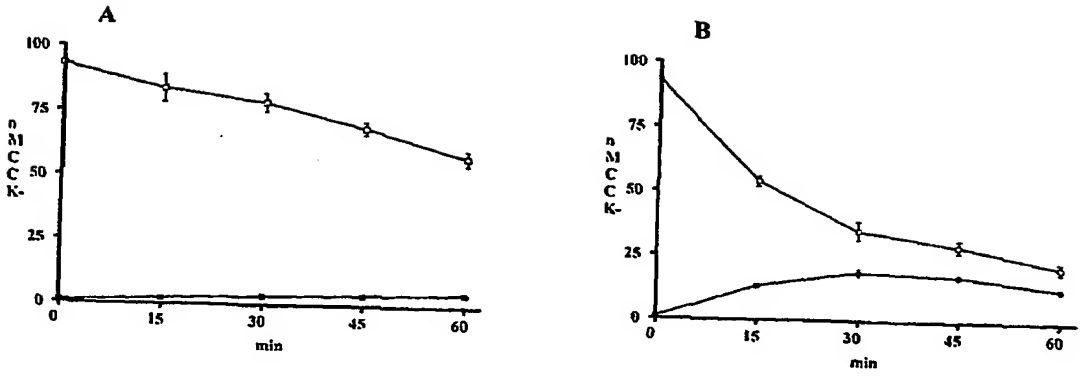

Figure 7

Modtaget PVS
20 SEP. 2002

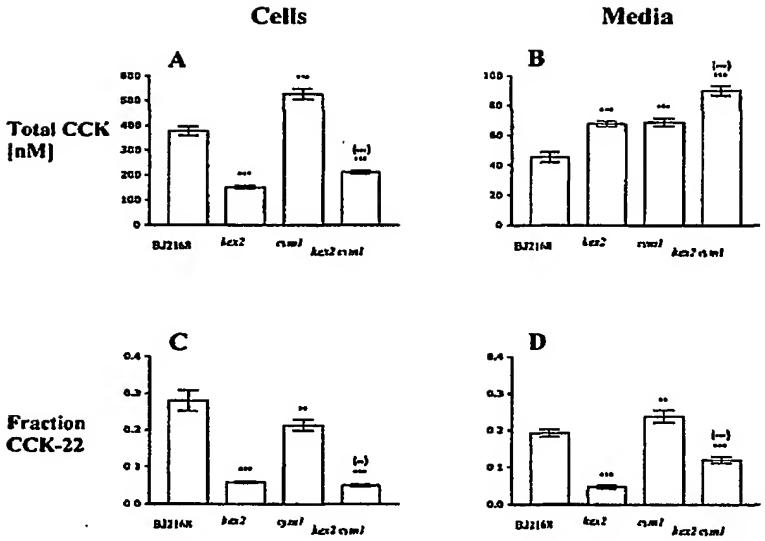

Figure 8

Modtaget PVS
20 SEP. 2002

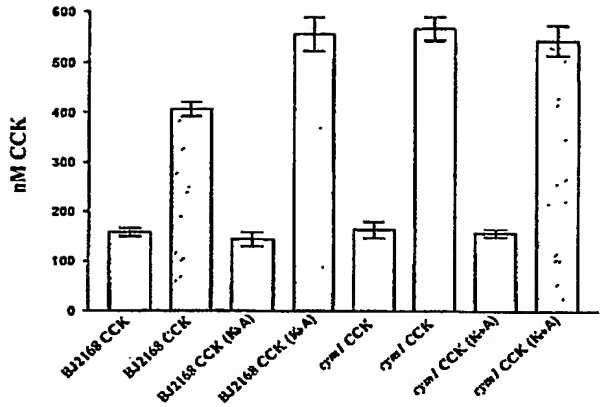

Figure 9

Modtaget PVS
20 SEP. 2002

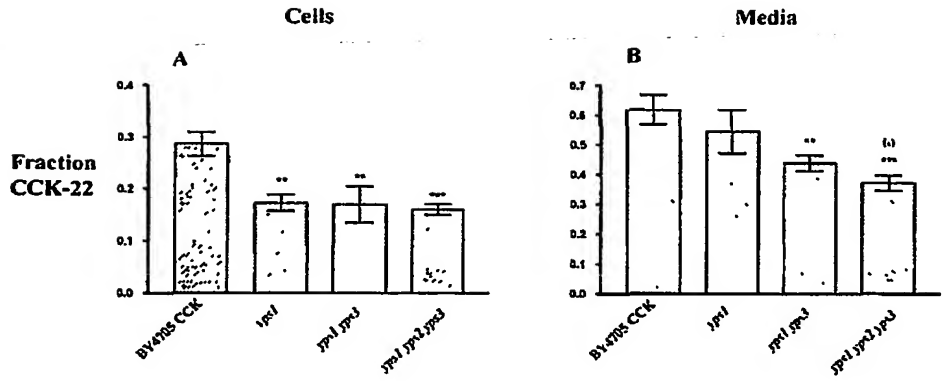

Figure 10

Modtaget PVS
20 SEP. 2002

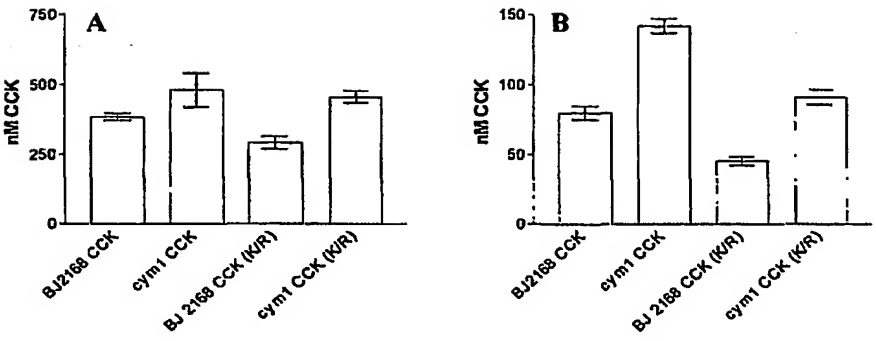

Figure 11

ProCCK	Molecular mass		Strain	SEQ ID NO:
CCK-61 CCK-58 CCK-39 CCK-22 CCK-8				
PreproMfα1p //.SGLQRAEEAPRRQLRVSQRTDGESRAHLGALLARYIQQARKAPSGRMSIVKNLQNLDPSHRISDRDYMGWMDFGRRSAEEYEYPS				43
QLRVSQRTDGESRAHLGALLAR	2433.3	2433.5	B	44
VSQRTDGESRAHLGALLAR	2036.1	2036.1	B	45
YIQQARKAPSGRMSIVKNLQNLDPSHRISDRDYMGWMDFGRRSAEEYEYPS	6051.9*	6051.6*	B	46
YIQQARKAPSGRMSIVK	1932.1	1932.2	A & B	47
YIQQARKAPSGRMSIV	1805.2	1805.0	A	48
NLQNLDPSHRISD	1509.0	1508.7	A	49
NLQNLDPSHRISDRDYMGWMDFG	2766.2	2766.1	A	50
NLQNLDPSHRISDRDYMGWMDFGRRSAEEYEYPS	4133.8	4133.9	A & B	51
DYMGWMDFGRRSAEEYEYPS	2488.0	2488.1	A	52

Figure 12

Mogtaget PVS
20 SEP. 2002

Modtaget PVS
20 SEP. 2002

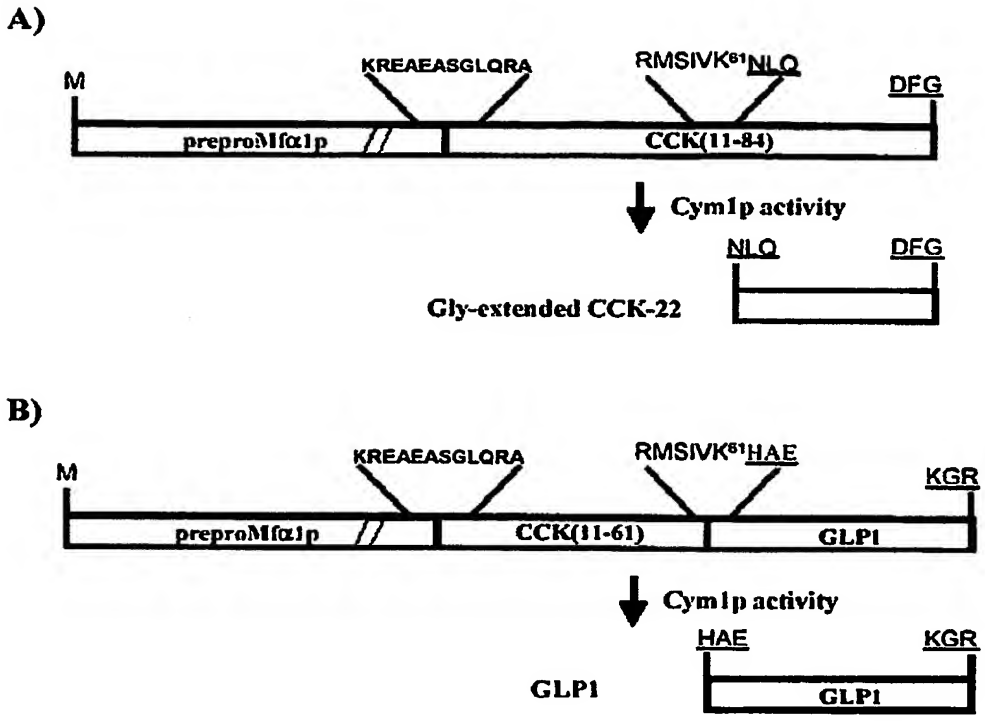

Figure 13